Exploring Archaeal Communities And Genomes Across Five Deep-Sea Brine Lakes Of The Red Sea With A Focus On Methanogens

Dissertation by

Yue Guan

In Partial Fulfillment of the Requirements

For the Degree of

Doctor of Philosophy

King Abdullah University of Science and Technology

Thuwal, Kingdom of Saudi Arabia

© *December 2015*

Yue Guan

All Rights Reserved

EXAMINATION COMMITTEE APPROVALS FORM

The dissertation of Yue Guan is approved by the examination committee.

Prof. Ulrich Stingl
King Abdullah University of Science and Technology
Committee Chairperson

Prof. Timothy Ravasi
King Abdullah University of Science and Technology

Prof. Pascal Saikaly
King Abdullah University of Science and Technology

Prof. James G. Ferry
Pennsylvania State University

ABSTRACT

Exploring Archaeal Communities And Genomes Across Five Deep-Sea Brine Lakes Of The Red Sea With A Focus On Methanogens

Yue Guan

The deep-sea hypersaline lakes in the Red Sea are among the most challenging, extreme, and unusual environments on the planet Earth. Despite their harshness to life, they are inhabited by diverse and novel members of prokaryotes. Methanogenesis was proposed as one of the main metabolic pathways that drive microbial colonization in similar habitats. However, not much is known about the identities of the methane-producing microbes in the Red Sea, let alone the way in which they could adapt to such poly extreme environments. Combining a range of microbial community assessment, cultivation and omics (genomics, transcriptomics, and single amplified genomics) approaches, this dissertation seeks to fill these gaps in our knowledge by studying archaeal composition, particularly methanogens, their genomic capacities and transcriptomic characteristics in order to elucidate their diversity, function, and adaptation to the deep-sea brines of the Red Sea. Although typical methanogens are not abundant in the samples collected from brine pool habitats of the Red Sea, the pilot cultivation experiment has revealed novel halophilic methanogenic species of the domain Archaea. Their physiological traits as well as their genomic and transcriptomic features unveil an interesting genetic and functional adaptive

capacity that allows them to thrive in the unique deep-sea hypersaline environments in the Red Sea.

ACKNOWLEDGEMENTS

I would to express my sincere appreciation and thanks to Professor Dr. Ulrich Stingl for being an incredibly supportive, patient, and inspiring supervisor and giving me the opportunity to be part of the Deep-Sea Brine project. For your guidance and support throughout the 4-year course of my doctoral research, I would like to thank you. I would also like to thank Professor Dr. Timothy Ravasi, Professor Dr. Pascal Saikaly, and Professor Dr. James G. Ferry for serving as my committee members. Your constructive criticism and suggestions improved the quality of my work substantially and helped make me a better scientist. I am also very grateful to Professor Dr. James Ferry for mentoring me on how to work with anaerobic cultures and sharing his vast knowledge of anaerobic metabolisms.

My appreciation also goes to my friends and colleagues in the Stingl Lab, the Red Sea Research Center, the Computational Bioscience Research Center, the Biosciences Core Lab, the Coastal and Marine Resources Core Lab and the Imaging Core Lab for making my doctoral research KAUST a memorable learning experience. My special thanks goes to David Ngugi, Vinu Manikandan, John Pearman, Shwen Ho, Craig Michell, Mohamed Haroon, Karie Holtermann, André Antunes, Sebastian Baumgarten, and Sylvain Guillot. I also want to extend my gratitude to the R/V Aegaeo and crew, which has made our expeditions to the deep-sea brines of the Red Sea successful, and to King Abdullah University of Science and Technology and the SEDCO Research Excellence Award for funding this the study.

Finally, my heartfelt gratitude is extended to my family for their encouragement and support. I would also like to thank Prof. David Keyes and Wendy Keyes for being amazing mentors and being family to me in Saudi Arabia, so far away from my native home. You have exemplified literally everything I dream to achieve. Your inspirations in life, science, and music have helped me become a more mature and happier person. Special thanks go to my dearest friends in KAUST, especially Jens Schneider, Greg Wickham, Jill Pagels, Shwen Ho, Christiane Hoppe-Jones, John Pearman, Ian Shore, Bilel Hamzaoui, Jelena Bajic, my Mexican and French friends in KAUST for everything; also thanks must go to Antonio Dvorak for composing his exceptional work of the 16 numbers of Slavonic Dances. Studying and performing them have greatly enriched my life while perusing this doctoral degree.

Last but not least, I would like to thank HRH King Abdullah. Without his vision and KAUST, I could not have ever imagined to be given the precious opportunity to explore the deep-sea brine lakes of the Red Sea and the unique culture of Saudi Arabia and the Middle East region.

I would like to dedicate this dissertation to Professor David E. Keyes.

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

AAI	Average amino acid identity
ANI	Average nucleotide identity
BLAST	Basic local alignment search tool
BSI	Brine-seawater interface
COGs	Cluster of Orthologous Groups
CRISPR	Clustered regularly interspaced short palindromic repeats
Cryo-SEM	Cyro-scanning electron microscopy
DHABs	Deep hypersaline anoxic basins
DMA	Dimethylamine
DNA	Deoxyribonucleic acid
DSMZ	German Collection of Microorganisms and Cell Cultures
INDIGO	INtegrated Data Warehouse of MIcrobial GenOmes
KAUST	King Abdullah University of Science and Technology
LCL	Lower convective layer
MCR	Methyl-coenzyme M reductase
Mhb.	*Methanohalobium*
Mhp.	*Methanohalophilus*
MMA	Monomethylamine
MSBL1	Mediterranean Sea Brine Lakes group 1
OTU	Operational taxonomic unit
PCR	Polymerase chain reaction
Pyl	Pyrrolysine
qPCR	Quantitative polymerase chain reaction
RNA	Ribonucleic acid
TCA cycle	Tricyclic acid cycle
TEM	Transmission electron microscopy
TMA	Trimethylamine
UCL	Upper convective layer

LIST OF FIGURES

LIST OF TABLES

INTRODUCTION

1. Archaea and domains of life

Prior to 1977, archaea were considered as bacteria because of their prokaryotic appearance. This changed when Dr. Carl Woese and colleagues initially proposed the concept of three domains of life by proposing “archaebacteria ('ancient' bacteria), containing methanogenic bacteria” as an additional domain of life (Woese and Fox, 1977). It took another 13 years before Archaea was formally recognized as the third domain of cellular life besides Bacteria and Eukarya (Woese *et al.*, 1990).

As they appeared as a sister lineage in Woese’s ‘universal tree’, Archaea and Eukarya have been hypothesized to share a common ancestor (Woese *et al.*, 1990). Archaea differ from bacteria in aspects, amongst others, such as membrane lipids, ribosomes, initiator tRNA, RNA polymerases, the requirement of transcription factors. (Madigan *et al.*, 2008). In 1996, after the first archaeal genome—*Methanocaldococcus jannaschii* (formerly known as *Methanococcus jannaschii)*—had been sequenced, the genes involved in DNA replication, transcription, and translation in this genome were found to be closer to their counterparts in eukaryotes than to those in other bacteria (Bult *et al.*, 1996).

Originally and for quite some time, the domain Archaea was considered to comprise only two phyla: the Crenarchaeota and the Euryarchaeota. In the recent few decades, our knowledge of this new domain has witnessed unparalleled growth. The Archaea are currently classified into the following diverse phyla/superphyla: Euryarchaeota, the so-called **TACK** superphylum consisting of:

Thaumarchaeota (so far identified as chemolithoautotrophic ammonia-oxidizers), **A**igarchaeota (with the main representative *Caldiarchaeum subterraneum*), **C**renarchaeota, **K**orarchaeota (so far only found in high temperature hydrothermal environments), Bathyarchaeota (Meng *et al.*, 2014) and Geoarchaeota ('novel archaeal group 1 (NAG1)') (Kozubal *et al.*, 2013) (were suggested within the TACK superphylum), Lokiarchaeota (Spang *et al.*, 2015), the **DPANN** superphylum containing **D**iapherotrites (pMC2A384), **P**arvarchaeota, **A**enigmarchaeota (DSEG) and **N**anoarchaeota (with the representative *Nanoarchaeum equitans* to be an obligate symbiont on the archaeon *Ignicoccus*), **N**anohaloarchaeota, Micrarchaeota, Woesearchaeota and Pacearchaeota (Garrity *et al.*, 2004; Huber *et al.*, 2002; Brochier-Armanet *et al.*, 2008; Barns *et al.*, 1996; Rinke *et al.*, 2013; Narasingarao *et al.*, 2012; Castelle *et al.*, 2015).

The first described archaea were from hydrothermal and hypersaline environments. Consequently, they were initially regarded as extremophiles living in harsh environments (e.g. hot, acidic, salty) (Lopez-Garcia *et al.*, 2015), but they have since been found colonizing an extraordinary range of habitats, including more moderate environments and constitute a substantial fraction of the microbiota on our planet. In mankind's brief period of investigative history into the domain Archaea, we have already been amazed by many of their remarkable features and critical functions.

2. Methane-producing members in the domain Archaea (methanogens)

Methane and methanogenesis

Methane has an important role in the global carbon cycle (Figure 1) (Thauer *et al.*, 2008). Contemporarily, methane is the most abundant hydrocarbon trace gas in the atmosphere and is a potent greenhouse gas, having a global warming potential 21 times that of carbon dioxide (CO_2) (Canfiel *et al.*, 2005). The current global methane emission is about 500-600 Tg methane (CH_4) per year (Lowe, 2006) due to the production of methane through the thermochemical or biological breakdown of organic matter being greater than the consumption of methane through the action of methanotrophic organisms.

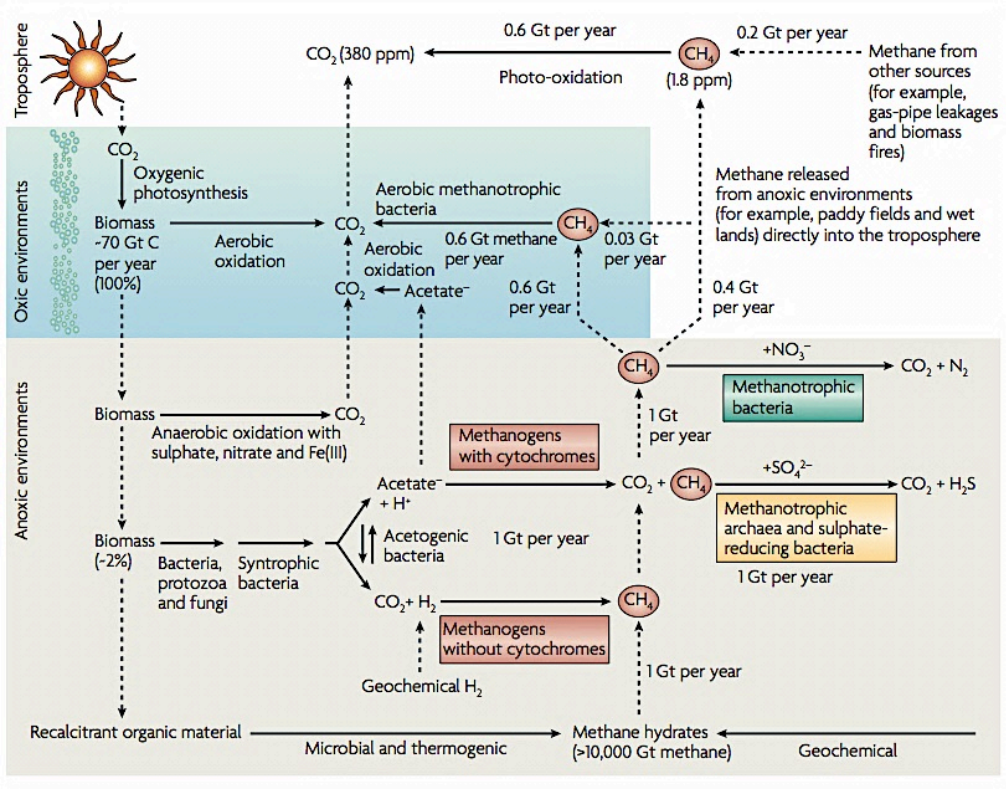


Figure 1. Methane as an intermediate in the global carbon cycle (from Thauer *et al.*, 2008).

About 74% of the emitted methane is produced globally by methanogenesis (Conrad, 2009; Whitman *et al.*, 2006). Approximately 2% of the net primary production (NPP) (70 giga-tonnes of carbon per year) that is fixed into biomass per annum by photosynthesis ends up recycled primarily through methane in anoxic environments (Thauer *et al.*, 2008).

Methanogenesis represents one of the earliest and ancient biological processes on Earth. The evidence of microbial methane production in ancient microbial ecosystems has now emerged from 3.5-billion-year-old rocks (Ueno *et al.*, 2006). Thus methanogenesis may have played a major role in the evolution of the Earth's atmosphere (Liu *et al.*, 2012; Ueno *et al.*, 2006).

Methanogens are the only kind of microorganisms known to produce methane as the end product of their anaerobic respiration. They are present in almost every conceivable anaerobic environment (Ferry, 1994). Methanogens are abundant in habitats where electron acceptors such as O_2, NO^{3-}, Fe^{3-}, and SO_4^{2-} are less abundant due to the lack of competition from the sulfate-reducing bacteria, denitrifying bacteria, and iron-reducing bacteria. The distribution of known species of methanogens in natural environments is also dependent of their individual adaptation to various temperatures, pH and salinity ranges. They can grow at a broad ranges of temperatures (from 4°C to 110°C), pH (from acidic to alkalinic), and salinities (from freshwater to marine water to brine) (Ferry, 1994). The most important natural sources of methane are gases released from wetland soils, lake sediments and the digestive tracts of termites (Canfiel *et al.*, 2005).

Halophilic methanogens also contribute significantly to carbon mineralization in marine and hypersaline environments (Spring *et al.*, 2010).

Substrates and pathways for methanogenesis

Methanogens use a restricted collection of simple carbon compounds. The substrate specification of each genus is described in Table 1. Based on these different substrates, methanogens can further be divided into three groups: hydrogenotrophic, aceticlastic and methylotrophic methanogens (Garcia *et al.*, 2000; Liu and Whitman, 2008; Ferry, 2001).

Table 1. Methanogenesis reactions, free energies and typical organisms (Hedderich and Whitman, 2006; Liu and Whitman, 2008).

Reaction	$\Delta G^{\circ\prime}$ [a] (kJ/mol CH_4)	Organisms
I. CO_2-type		
$4\ H_2 + CO_2 \rightarrow CH_4 + 2\ H_2O$	−135	Most methanogens
$4\ HCOOH \rightarrow CH_4 + 3\ CO_2 + 2\ H_2O$	−130	Many hydrogenotrophic methanogens
CO_2 + 4 isopropanol → CH_4 + 4 acetone + 2 H_2O	−37	Some hydrogenotrophic methanogens
$4\ CO + 2H_2O \rightarrow CH_4 + 3\ CO_2$	−196	*Methanothermobacter* and *Methanosarcina*
II. Methylated C1 compounds		
$4\ CH_3OH \rightarrow 3\ CH_4 + CO_2 + 2\ H_2O$	−105	*Methanosarcina* and other methylotrophic methanogens
$CH_3OH + H_2 \rightarrow CH_4 + H_2O$	−113	*Methanomicrococcus blatticola* and *Methanosphaera*
$2\ (CH_3)_2\text{-}S + 2\ H_2O \rightarrow 3\ CH_4 + CO_2 + 2\ H_2S$	−49	Some methylotrophic methanogens
$4\ CH_3\text{-}NH_2 + 2\ H_2O \rightarrow 3\ CH_4 + CO_2 + 4\ NH_3$	−75	Some methylotrophic methanogens
$2\ (CH_3)_2\text{-}NH + 2\ H_2O \rightarrow 3\ CH_4 + CO_2 + 2\ NH_3$	−73	Some methylotrophic methanogens
$4\ (CH_3)_3\text{-}N + 6\ H_2O \rightarrow 9\ CH_4 + 3\ CO_2 + 4\ NH_3$	−74	Some methylotrophic methanogens
$4\ CH_3NH_3Cl + 2\ H_2O \rightarrow 3\ CH_4 + CO_2 + 4\ NH_4Cl$	−74	Some methylotrophic methanogens
III. Acetate		
$CH_3COOH \rightarrow CH_4 + CO_2$	−33	*Methanosarcina* and *Methanosaeta*

Hydrogenotrophic methanogens (species in the orders *Methanobacteriales* [with the exception of *Methanosphaera*], *Methanococcales*, *Methanomicrobiales*, Methanopyrales and Methanocellales) produce methane via the reduction of CO_2 to a methyl group with electrons derived from the oxidation of electron donors (primarily H_2 or formate). Methylotrophic methanogens (species in the order of *Methanosarcinales* and the 7th order of *Methanomassiliicoccales*), utilize methyl-containing C–1 compounds as methanol, various methylamines, or

dimethylsulfides to produce methane. In *Methanosarcinales*, a molecule of substrate is demethylated to provide the methyl group and another molecule is oxidized to provide electrons for reduction of CoM-S-S-CoB (Ferry, 2001). Methanogens belonging to *Methanomassiliicoccales* lack the majority of steps to oxidize methyl-groups. Instead, their methylotrophy is dependent on H_2 (Borrel *et al.*, 2014). Although rarely used, betaine, choline methylated sulfur compounds (dimethylsulfide), and methanethiol could also serve as substrates in certain methylotrophic methanogens (Watkins *et al.*, 2012; 2014; Mathrani *et al.*, 1988; Finster *et al.*, 1992). Aceticlastic methanogens (also called acetotrophic methanogens) cleave acetate to provide a methyl group for CH_4 genesis and provide a carbonyl group for oxidation to CO_2. The methane generated by the acetoclastic pathway makes up at least two-thirds of CH_4 in the biosphere. However, *Methanosarcina* and *Methanothrix* (*Methanosaeta*) are the only two genera known to date that can grow and produce CH_4 from acetate fermentation (Ferry, 2010). *Methanosaeta pelagica* is the only *Methanosaeta* species isolated from marine environments and grows optimally at a salinity of 0.28M (1.6%) but can grow up to a salinity of 0.80M (4.7%) (Mori *et al.*, 2012), which is in contrast to other *Methanothrix* species which can not tolerate elevated concentrations of salt.

Unique, although partially overlapping, pathways are utilized to metabolize various substrates. Only three types of methanogenic pathways are known: CO_2-reduction, methyl-group reduction (H_2-independent or H_2-dependent), and the aceticlastic reaction (Liu and Whitman, 2008; Ferry, 2010) In hypersaline environments, factors such as redox potential, permanency of anaerobic

conditions, and the concentration of other terminal electron acceptors determine methanogenesis. Hydrogenotrophic and acetoclastic methanogens are more vulnerable compared to methylotrophic methanogens due to their lower energy yield (Lazar *et al.*, 2011; McGenity, 2010). However, in most freshwater environments, the acetoclastic methanogens are responsible for approximately two-thirds of CH_4 production with most of the remaining one-third produced by CO_2 reducing methanogens (Ferry, 2011).

Phylogeny and taxonomy of methanogenic archaea

Methane-producing microorganisms are primarily affiliated with the phylum Euryarchaeota. Based on the phylogenetic analysis using marker gene and ribosomal proteins sequences, the methanogens are currently classified into seven orders, *Methanopyrales*, *Methanococcales*, *Methanobacteriales*, *Methanomicrobiales*, *Methanosarcinales*, *Methanocellales* (Sakai *et al.*, 2008) and *Methanomassiliicoccales* (Paul *et al.*, 2012; Borrel *et al.*, 2013).

Based on the information collected from the book *The Prokaryotes – Other Major Lineages of Bacteria and the Archaea* (Rosenberg *et al.*, 2014) (information for *Methanomassiliicoccus* is collected from Dridi et al., 2012 (Dridi *et al.*, 2012) and Brugere *et al.*, 2014 (Brugère *et al.*, 2014), I provide a detailed summary of the current taxonomy of all seven methanogenic archaea up to the genus level with their substrate utilization and salinity range in Table 2.

Table 2. Methanogen taxonomy: order, family, and genus, with a note on substrate utilization and salinity range.

Order	Family	Genus	Substrates	Salinity range (%)
Methanobacteriales	*Methanobacteriaceae*	*Methanobacterium*	H_2/CO_2, formate	0 – 7
		Methanothermobacter	H_2/CO_2, formate	0.0025 – 5
		Methanobrevibacter	H_2/CO_2, formate	up to seawater
		Methanosphaera	Methanol + H2	NR
	Methanothermaceae	*Methanothermus*	H_2/CO_2	NR
Methanococcales	*Methanococcaceae*	*Methanococcus*	H_2/CO_2, formate, pyruvate + CO_2	0.3 - 5
		Methanothermococcus	H_2/CO_2, formate	0.6 - 9.6
	Methanocaldococcaceae	*Methanocaldococcus*	H_2/CO_2	0.5 - 5.6
		Methanotorris	H_2/CO_2 or with formate	0.4 - 6.0
Methanomicrobiales	*Methanomicrobiaceae*	*Methanomicrobium*	H_2/CO_2, formate	NR
		Methanoculleus	H_2/CO_2, formate, alcohols	0 - 9.4
		Methanolacinia	H_2/CO_2, formate, alcohols	0 - 4.7
		Methanoplanus	H_2/CO_2, formate, alcohols	1.0 - 5
		Methanogenium	H_2/CO_2, formate, alcohols	1.5 - 7.3
		Methanofollis	H_2/CO_2, formate, alcohols	0 - 7
	Methanocorpusculaceae	*Methanocorpusculum*	H_2/CO_2, formate, alcohols	0 - 4.7
	Methanospirillaceae	*Methanospirillium*	H_2/CO_2, formate	0 - 0.6
	Methanocalculaceae	*Methanocalculus*	H_2/CO_2, formate	0 – 12.5
	Methanoregulaceae	*Methanolinea*	H_2/CO_2, formate	0 - 2.5
		Methanoregula	H_2/CO_2, formate	0 - 1
		Methanosphaerula	H_2/CO_2, formate	0 - 0.5
Methanopyrales	*Methanopyraceae*	*Methanopyrus*	H_2/CO_2	0.2 - 4
Methanocellales	*Methanocellaceae*	*Methanocella*	H_2/CO_2, formate	0 - 2
Methanosarcinales	*Methanosarcinaceae*	*Methanosarcina*	H_2/CO_2 or CO, acetate, methanol, methylamines, dimethylsulfide	0 - 6
		Methanococcoides	Methanol, methylamines	0.58 - 4.1
		Methanohalophilus	Methanol, methylamines	1.8 - 20.5
		Methanohalobium	Methanol, methylamines	15.2 - 29.8
		Methanosalsum	Methanol, methylamines, dimethylsulfide	1.17 - 12.3
		Methanomethylovorans	Methanol, methylamines, dimethylsulfide, methanethiol	0 - 1.8
		Methanolobus	Methanol, methylamines, methyl sulfides	0.35 - 11.7
		Methanomicrococcus	Methanol + H_2, methylamines + H_2	0 - 1.8
		Halomethanococcus	Methanol, methylamines	10.5 - 21.0
	Methermioccaceae	*Methermicoccus*	Methanol, methylamine, trimethylamine	1.2 - 6.4
	Methanotrichaceae (Methanosaetaceae)	*Methanothrix (Methanosaeta)*	Acetate	0 - 0.6
		"Methanosaeta pelagica" sp. nov	Acetate	1.2 - 4.7
Methanomassiliicoccales	*Methanomassiliicoccaceae*	*Methanomassiliicoccus*	Methanol + H2, methylamines + H2	0.1 - 1.0

Methanogens in hypersaline environments

Halophilic methanogens contribute significantly to carbon mineralization in marine and hypersaline environments. Methanogenesis has been studied in a number of hypersaline environments, such as the Mediterranean Sea, the Great Salt Lake and Shark Bay (McGenity, 2010; Zhilina, 1986). Pure cultures of methanogens growing on different substrates have been shown to have different tolerance of salinity (Oren, 2002) (see also Table 2 and Table 3). Methanogens using the acetoclastic pathway showed the least salt tolerance (up to 6%). *Methanocalculus halotolerans*, the most salt-tolerant methanogen that uses H_2/CO_2 for energy generation, can grow in cultivation media with up to 12.5 % NaCl (Ollivier *et al.*, 1998). Greater tolerance of high salt concentrations has been observed in various methylotrophic methanogens. For instance, the most salt tolerant methanogenic species known to date is *Methanohalobium evestigatum* with an optimum growth salinity of 25% and a maximum growth salinity of 30% (Zhilina and Zavarzin, 1987) (see also Table 3).

Table 3. The upper salt concentration limits of methanogens with different energy-yielding processes.

Methanogenesis process	Most halotolerant representative	Optimal NaCl (%)	Approx. upper salt limit in culture
From acetate	*Methanosarcina acetivorans*	~2%	6%
From H_2 and CO_2	*Methanocalculus halotolerans*	5%	12.5%
From methylated amines or methanol	*Methanohalobium evestigatum*	25%	30%

These halophilic obligately methylotrophic methanogens are taxonomically affiliated with several genera within the *Methanosarcinaceae* family with the least tolerant belonging to the genera *Methanolobus* and *Methanococcoides*, moderately halophilic methylotrophic species belong to the *Methanohalophilus* and *Methanosalsum* genera, and extreme halophiles being classified in the genus *Methanohalobium*. The main cultivated and taxonomically described obligately halophilic species are described in Table 4.

Table 4. Cultivated, taxonomically described halophilic methylotrophic methanogens. (Based on data from Euzeby (http://www.bacterio.cict.fr/m/methanocalculus.html), Joint Genome Institute (http://www.jgi.doe.gov/) and McGenity (McGenity, 2010).)

Species	Habitat	NaCl range (%)	NaCl opt (%)	Carbon sources	Original publication	Genome
Methanococcoides methylutens	The Sumner branch of Scripps Canyon, USA, sediment	1.4 - 6	2.47	Methanol, methylamines	Sowers et al., 1983	Guan et al., 2014
Methanococcoides burtonii	Ace Lake, Antarctica			Methanol, methylamines	Franzmann et al., 1992	Allen et al., 2009
Methanohalophilus mahii (DSM 5219)	Great Salt Lake, sediment	3 - 20	6 - 15	Methanol, methylamines	Paterek and Smith, 1988	Spring 2010
Methanohalophilus portucalensis (DSM 7471)	Saltern sediments, Portugal	2 - 25	3 - 12	Methanol, methylamines	Boone et al., 1993	
Methanohalophilus halophilus (DSM 3094)	Shark Bay, Australia, microbial mat	2 - 15	7 - 9	Methanol, methylamines	Zhilina 1984	
Methanohalophilus zhilinae (DSM 4017)	Bosa Lake (Wadi Natrun)	1 - 12	4	Methanol, methylamines, DMS	Mathrani et al., 1988	Published sequence only
Methanohalobium evestigatum (DSM 3721)	Sivash Lake, microbial mat, Crimia	15 - 30	25	Methanol, methylamines	Zhilina and Zavarzin, 1988	Published sequence only

In hypersaline environments, methylotrophic methanogens are the most commonly observed methanogens. The C-1 methylated compounds are continuously provided by the degradation of osmolytes (such as glycine

betaine) or membrane compounds (such as choline) (see also Figure 2) (McGenity, 2010).

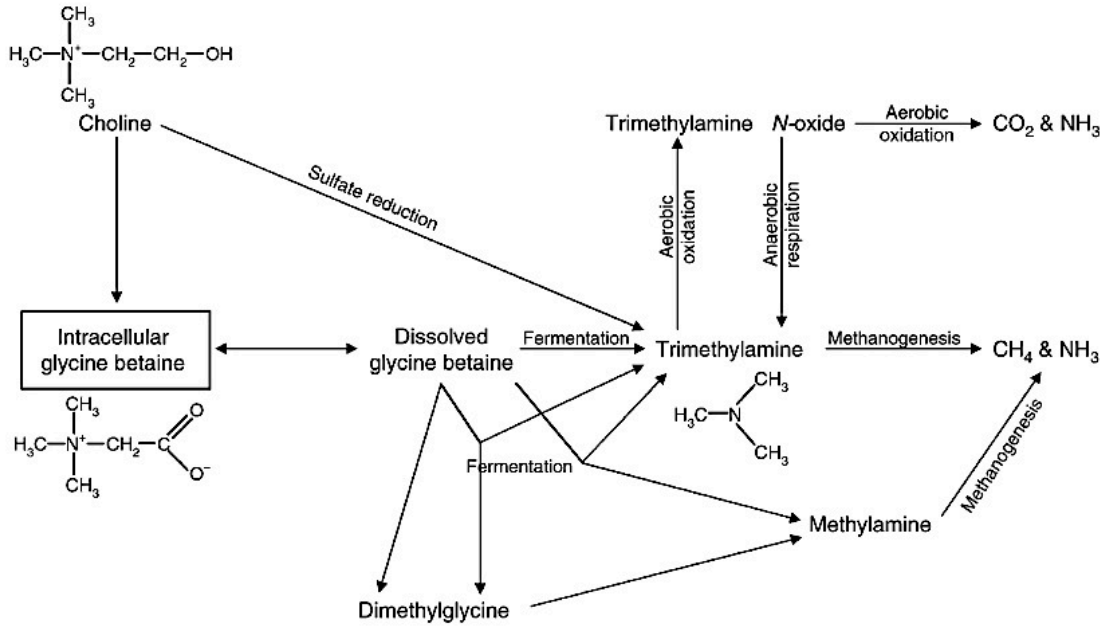


Figure 2. Central importance of trimethylamine and main pathways for methanogenesis in hypersaline environments (McGenity, 2010).

Unusual amino acid in methanogens

In addition to the 20 canonical amino acids methanogens have been observed to use other amino acids. For example some members of CO_2 fixation methanogens can utilize the proteinogenic amino acid selenocysteine (Rother and Krzycki, 2010) whilst methylotrophic methanogens have been shown to encode the proteinogenic amino acid pyrrolysine (Hao *et al.*, 2002). Both of these unusual amino acids are incorporated into proteins at positions, which usually result in a termination codon UGA or UAG, respectively on their mRNA translation template.

Selenoprotein-containing enzymes have bee found in many key enzymes, such as formate dehydrogenase, formyl-methanofuran dehydrogenase, F_{420}-reducing hydrogenase, F_{420}-nonreducing hydrogenase, and heterodisulfide reductase in hydrogenotrophic methanogenesis of *Methanococcus*, *Methanopyrus*, and *Methanocaldococcus* (Vorholt *et al.*, 1997; Jones *et al.*, 1979; Jones and Stadtman, 1981; Halboth and Klein, 1992; Sorgenfrei *et al.*, 1993). Selenocysteine also has been found to be incorporated in selenophosphate synthetase that is directly involved in selenocysteine biosynthesis of selenophosphate, the selenium donor used to synthesize selenocysteine (Stock *et al.*, 2010).

Pyrrolysine was first identified at the active site of methylamine methyltransferases, the enzyme that initiates in methanogenesis through the methyl group transfer to the cobalt ion of a cognate corrinoid protein in methylotrophic methanogens (Hao *et al.*, 2002; Krzycki, 2004). To date, no exception has been observed and all methanogen mono-, di- and trimethylamine methyltransferase genes have the common characteristic of an in-frame amber codon. Pyrrolysine could also appear in other proteins in methanogens as a result of neutral evolution (Heinemann *et al.*, 2009). Unlike decoding of selenocysteine, which has been found in all three domains of life, only a few organisms have the ability to decode pyrrolysine according to sequenced genome data. These are all known methyl-compound utilizing methanogens as well as some bacteria the majority of which are affiliated with *Clostridia* and *Deltaproteobacteria* (Prat *et al.*, 2012; Gaston *et al.*, 2011).

3. Deep-sea anoxic hypersaline brines of the Red Sea and microbial explorations

Geographic locations

Hypersaline systems are harsh environments that are defined as those with much higher salt concentrations than the surrounding seawater, often close to or exceeding salt saturation. Deep-sea anoxic hypersaline brines, which are also referred to as deep hypersaline anoxic basins, are considered to be one of most remote, challenging and extreme environments on Earth (Antunes *et al.*, 2011). They have been found under a wide range of geodynamic situations in the Red Sea, the Mediterranean Sea, and the Gulf of Mexico.

In the six decades since the discovery of the Albatross Deep in 1948 (Meurant, 1976), intensive oceanographic investigations have been undertaken to discover new brine deeps in the Red Sea and elucidate the unusual geological and physicochemical characters of the brines of the Red Sea. To date, there are approximately 25 deep-sea brines in the Red Sea (Figure 3). They are located along the rift valley (central graben axis between 19°N and 27°N) in the center of the Red Sea where the Asian Plate and African Plate diverge (Schmidt *et al.*, 2015). Their depths range from 1,196 meters (observed in the Vema Deep) to 2,850 meters (observed in the Southwest Basin of the Suakin Deep) below the sea surface (Backer and Schoell, 1972). Their thicknesses vary from 39 meter (Nereus Deep) to 322 meter (Port Sudan Deep) (Antunes *et al.*, 2011).

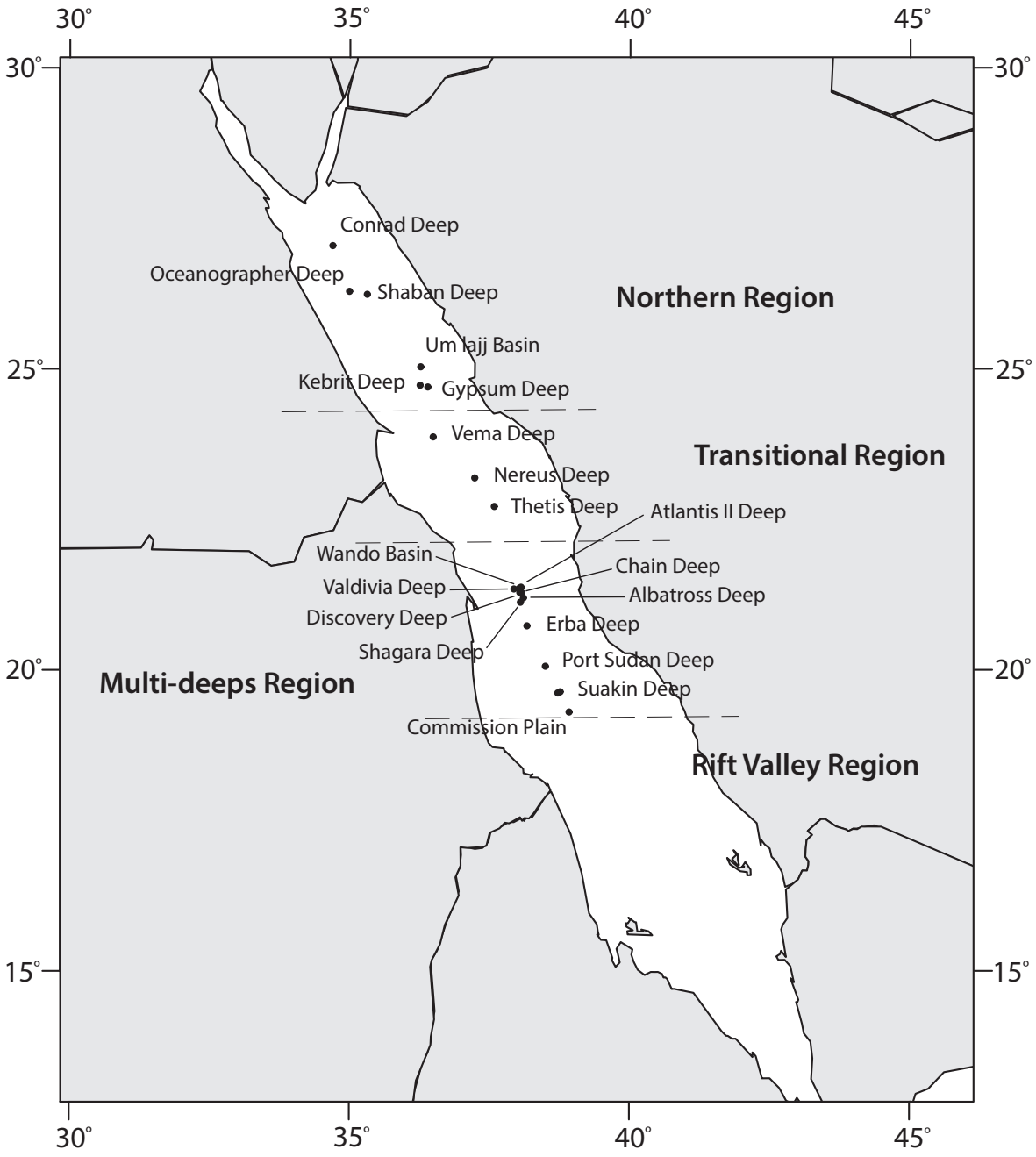


Figure 3. Geographic location of some major deeps and basins of the Red Sea. Coordinates of the individual brines that used for plotting are collectedly obtained from Backer et al., Edwards et al., and Cochran et al. ((Backer and Schoell, 1972; Edwards and Head, 1987; Cochran *et al.*, 1986).The axial zone of the Red Sea can be divided into four regions based on morphotectonic characteristics. From the north to south: the definition of the four regions of the red Sea: Northern Region, Transitional Region, Multi-deeps Region, and Rift Valley Region, according to Bonatii et al. (Bonatti, 1985)

Physical and geochemical features and classifications

There are considerable variations in the physical and geochemical parameters of these deep-sea brines of the Red Sea (Backer and Schoell, 1972; Antunes *et al.*, 2011). Among many other variables, they feature significant differences in salinities (from about 5% NaCl concentration in the brine-seawater interface of Atlantis II Deep to 26% NaCl concentration in the brine layer of Atlantis II Deep), in temperatures (from 22.8°C in Shaban Deep to 68.2°C in Atlantis II Deep), pH values (from 5.0 in Atlantis II Deep to 7.7 in Suakin Deep), H_2S concentrations (ranging from not detectable in the majority of the deeps to 24 mg/l in Kebrit Deep), CH_4 concentrations (from 3810 ng/l in Nereus Deep to 9.5 x 10^6 ng/l in Kebrit Deep). Some of the chemical concentrations are not significantly higher than the overlaying seawater. Many others, however, vary significantly between the brine and the seawater layer, e.g. the Fe concentration in Atlantis II Deep (99.89 mg/l) is almost 200 times of that in the seawater (0.50 mg/l).

Based on statistical cluster analyses of geochemical brine data (as explained in (Schmidt *et al.*, 2015) and the references therein), some of these brine-filled deeps could be further classified into the following two types. “Type I” brines with Kebrit Deep and Oceanographer Deep as examples are described as “collapse structure” deep formation. They are located within the Northern Region of the Red Sea and they mainly exhibit high salinity and H_2S concentration, but low pH values, low temperatures, and low trace metal concentrations. “Type II” brines with Atlantis II Deep as a representative are linked to “intrusion-/extrusion-related” deep formation. Many of the Type II

deeps are located within the Transitional Region and the Multi-deeps Region. They are thermally active and have low H_2S concentrations, but, in contrast, exhibit high element concentrations (e.g. Fe, Mn, Zn).

In addition, all these brines form characteristic brine-seawater interfaces due to the density gradient between seawater (lighter) and the brine body (heavier). Figure 4 depicts temperature gradients of two Red Sea brine deeps. Interfaces are immediately apparent from the steep increase of temperature between characteristic temperature “plateaus”. Additional distinctive features that are generally associated with brine-seawater interfaces, including sharp gradients of salinity, density, O_2, and pH. The brine/seawater interface also acts as an in situ particle trap for organic and inorganic materials from the seawater (Antunes *et al.*, 2011). Pronounced stratification of the brine layers is also observed in Atlantis II Deep comprising a pile of four convective layers with gradual and structured temperature and salinity increments (Balnc and Anschutz, 1995; Swift *et al.*, 2012), with the layers being destabilized by diffusive vertical heat flux and stabilized by salinity and densities (Antunes *et al.*, 2011). As a consequence of the interaction of the anoxic brines with the oxidizing, weakly alkaline Red Sea waters, deposition of heavy metals is favored resulting in large amounts of metalliferous sediments located at the bottom of the brines (Antunes *et al.*, 2011).

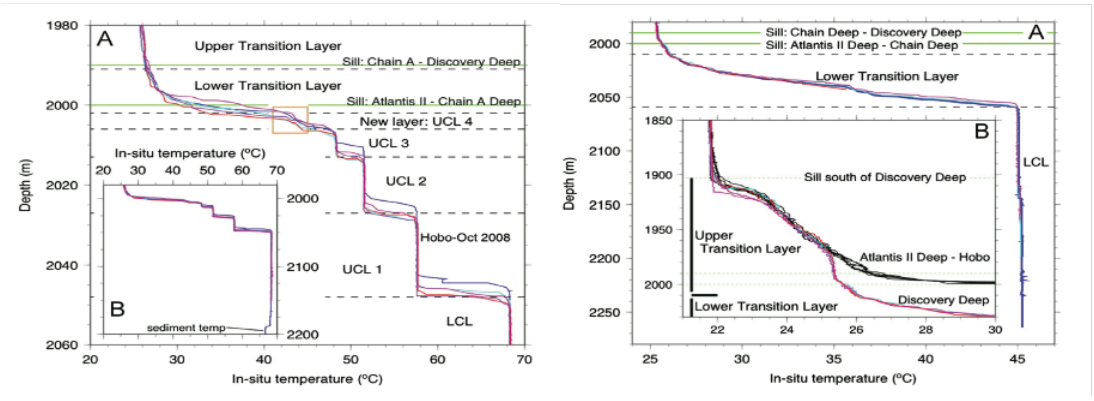


Figure 4. Vertical temperature gradients of the water column, brine-sea water interfaces and brines of the Red Sea. Left: temperature profile of the Atlantis II Deep. Right: temperature profile of the Discovery Deep (Swift *et al.*, 2012).

Microbial diversity of the deep-sea brine pools of the Red Sea

Due to the harshness of their conditions, deep-sea brines are considered inhospitable to life. The initial microbial study on the Red Sea brines has regarded them as "sterile" due to negative outcome in cultivation endeavors (Watson and Waterbury, 1969) and, before 1999, virtually not much is known about their microbial aspect. In light of modern microbiology and advances in sequencing technology, phylogenetic studies have greatly changed our view on the microbiology of the deep-sea brines of the Red Sea. Phylogenetic studies have been conducted on samples taken from sediment and the brine-seawater interface of Kebrit Deep (Eder *et al.*, 1999; 2001), as well as on samples from the brine-seawater interface of Shaban Deep (Eder *et al.*, 2002). In all of these studies, small-subunit rRNA genes were amplified by PCR from community genomic DNA extracted from the sampled environment and subsequently cloned, and sequenced. The presence of unique 16s rRNA gene sequences suggest the presence of novel high-salt and high-pressure communities in the Red Sea, such as the candidate division KB1, which branches in between the

Aquificales and the *Thermotogales*, as well as the putatively novel bacterial SB1 and archaeal SA1 and SA2 groups. Sequence comparisons showed that the majority of ribosome sequences retrieved had low similarity to previously cultivated archaea or bacteria, reflecting the novelty of microbial communities that thrive under the harsh environmental conditions in the Red-Sea brines.

Very few successful pioneer cultivation attempts from Red Sea deep-sea brines have been undertaken and reported during the period 1969–2010. Only a few bacterial and archaeal strains have been obtained. Described species were isolated from the Atlantis II Deep – *Flexistipes sinusarabici* gen. nov, sp. nov. (Fiala *et al.*, 1990); isolated from Shaban Deep — *Salinisphaera shabanensis* gen. nov., sp. nov. (Antunes *et al.*, 2003), *Marinobacter salsuginis* sp. nov. (Antunes *et al.*, 2007), *Halorhabdus tiamatea* sp. nov. (Antunes, Taborda, *et al.*, 2008), and *Haloplasma contractile* ord. nov., fam. nov., gen. nov. (Antunes, Rainey, *et al.*, 2008). No methanogenic cultures have ever been obtained from the deep-sea brines of the Red Sea.

Because of political and geographic reasons, expedition trips to the Red Sea have been separated by up to 10 years from one another (Swift *et al.*, 2012). With the establishment of King Abdullah University of Science and Technology (KAUST) in 2009, regular and more frequent expeditions have made it possible to explore more thoroughly these deep-sea brines of the Red Sea. Within only 4 years (between 2010 and 2013), the brine pools of the Red Sea have been sampled three times during joint cruises on the R/V Aegaeo staffed by researchers from KAUST, the American University of Cairo, and Hong Kong

University of Science and Technology. In this period, multiple approaches were combined to obtain a more in-depth characterization of the microbial community structure, phylogeny and predicted functionalities of the "microbial dark matter" of the brine environment of the Red Sea. To name a few, the aforementioned approaches included: the metagenomic approach (Wang *et al.*, 2013; Abdallah *et al.*, 2014), the pyrotag of 16S rRNA gene amplicon approach (Siam *et al.*, 2012), the single-cell genomics approach (Ngugi *et al.*, 2014), and the clone library approach of marker genes (Guan *et al.*, 2015; Wang *et al.*, 2013; Abdallah *et al.*, 2014).

Based on sequence similarities, the microbial communities thrive in the deep-sea brines of the Red Sea are very novel to what we have cultivated before. Notably, the sequences related to the ubiquitous brine-pool-associated group candidate division MSBL I (Mediterranean Sea Brine Lakes group 1) was also discovered in the Red Sea (Guan *et al.*, 2015; Ngugi *et al.*, 2014). Hitherto, MSBL1 was already known to be highly dominant in anoxic hypersaline ecosystems in Mediterranean deep brines (Borin *et al.*, 2009; La Cono *et al.*, 2011; Yakimov *et al.*, 2013). Few 16S rRNA sequences related to known methanogens were found in the Mediterranean deep-sea brines; however, CH_4 has been found in all four Mediterranean basins. Methanogenesis has been proposed as one of the primary pathways in the deep-sea brine environment (Borin *et al.*, 2009). MSBL1 has been suggested to contribute to methanogenesis at high salinity (van der Wielen, 2005). In addition, *Methanohalophilus*-like and *Methanomassiliicoccus*-like populations are also

reported in the investigated Mediterranean deep brines (La Cono *et al.*, 2011; Yakimov *et al.*, 2013; 2014).

4. Living with salt

All three of the primary domains have halophilic or halotolerant members but these are restricted in phylogenetic diversity. High salt concentrations are toxic for most organisms because it greatly reduces the availability of water thus life at high salt concentrations is energetically costly. As proposed by Oren in 2011 (Oren, 2011), the main determination of whether a certain type of microorganism can make a living at high salt are: 1) the amount of energy generated during its dissimilatory metabolism. Well-adapted halophilic microorganisms are commonly found to use light (generally not energy limited), aerobic respiration, denitrification and other highly exergonic dissimilatory processes as their energy source (with high negative $\Delta G^{\circ\prime}$) (Figure 5); 2) the mode of osmotic adaptation used. In the prokaryotic world, two fundamentally different strategies exist to enable microorganisms to cope with the osmotic stress inherent to the presence of high salt concentrations. A) The “salt-in” strategy – cells may maintain high intracellular salt concentrations, osmotically at least equivalent to the external concentrations. B) The “compatible-solute” strategy – organic compatible solutes are used by cells to balance the environmental osmotic pressure. Thus it allows cells to maintain low salt concentrations within their cytoplasm (Oren, 1999).

Thus, the current general consensus is that successful halophiles generally yield great amount of energy during their dissimilatory metabolism and

employ the mode of osmotic adaptation that requires less ATPs. A few halophilic-lifestyle exceptions are also exit among organisms that are obtaining only little energy and (some of them) could use more energetically costly mode of osmotic adaptation, such as *Natranaerobius*, *Desulfosalsimonas propionicica* and some members of anaerobic methane oxidizers (Oren, 2011).

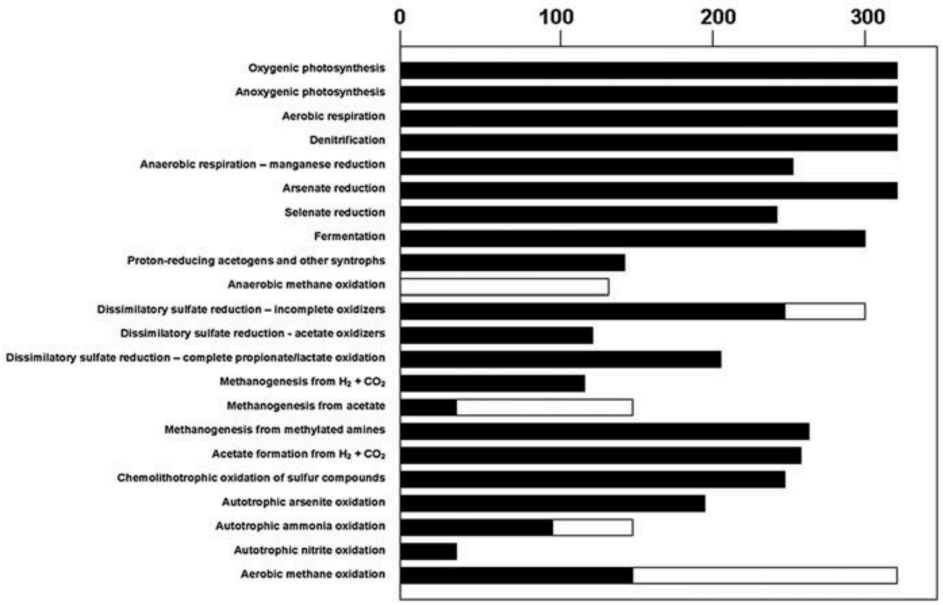


Figure 5. Approximate upper salt concentration limits for the occurrence of selected microbial processes (Oren, 2011).

5. Objectives and organization of chapters

Deep-sea brine pools of the Red Sea are among the most unique, polyextreme, deep, and least studied ecosystems on the planet Earth. The objectives of this dissertation are to study archaeal diversity, particularly methanogens, their genomic and transcriptomic characteristics, and how could they adapt in the deep-sea brines of the Red Sea.

In Chapter I, I present results obtained from intensively sampling five distinct brines, namely Erba Deep, Kebrit Deep, Nereus Deep, Discovery Deep, and Atlantis II Deep, in the Red Sea during the 3rd KAUST Red Sea Expedition in 2011. My colleagues and I studied the archaeal community diversity and phylogeny. We discovered that these brines harbor diverse archaeal lineages. Many of these lineages are without cultivated representatives. Archaeal members that could be related to typical methane-producers are not abundant in the sampled locations. We also found that many archaeal lineages are ubiquitously associated with the deep-sea brines, which indicates that they could be well adapted to the deep-sea brine ecosystem.

Data obtained were published in the main articles and supplementary materials of:

Guan Y, Hikmawan T, Antunes A, Ngugi DK, Stingl U. 2015. Diversity of methanogens and sulfate-reducing bacteria in the interfaces of five deep-sea anoxic brines of the Red Sea. *Research in Microbiology*. http://dx.doi.org/10.1016/j.resmic.2015.07.002.

Ngugi DK, Blom J, Alam I, Rashid M, Ba-Alawi W, Zhang G, Hikmawan T, Guan Y, Antunes A, Siam R, Dorry HE, Bajic V and Stingl U. 2015. Comparative genomics reveals adaptations of a halotolerant thaumarchaeon in the interfaces of brine pools in the Red Sea. *The ISME Journal*. 9, 396–411.

In Chapter II, I describe the cultivation and physiological characterization of a novel and first methane-producing isolate from the Erba Deep that is affiliated

with the extremely halophilic methanogen genus *Methanohalobium* and its genome. We performed a comparative genomics investigation of the Erba Deep isolate with its close and only cultivated relative *Methanohalobium evestigatum*. We furthermore conducted a comparative transcriptomic study of the Erba Deep isolate under different salinity conditions and examined its transcriptomic characteristics in details.

In a similar fashion, in Chapter III, I describe how we obtained a novel isolate from the Kebrit Deep that belongs to the moderately halophilic methanogen genus *Methanohalophilus*. The genomes of this novel isolate and three other members of this genus were sequenced for the first time and compared among each other in order to investigate the genomic differentiation among *Methanohalophilus* populations from varying hypersaline environment around the world.

For Chapter IV, we looked into single-amplified genomes from the deep-sea brines of the Red Sea in search of clues for methanogenic metabolism or key features that have previously mainly observed in methanogenic archaea. The presence of divergent pyrrolysine biosynthesis genes and enzymes that putatively incorporate pyrrolysine suggest potential codon expansion in members of the Mediterranean Sea Brine Lakes group 1 (MSBL1).

For each of the Chapters II to IV, a submission to peer-reviewed journals is under preparation.

The genome data of *Methanococcoides methylutens* used in some comparative analysis was first published in

Guan Y, Ngugi DK, Blom J, Ali S, Ferry JG, Stingl U. 2014. Draft genome sequence of an obligately methylotrophic methanogen, *Methanococcoides methylutens*, isolated from marine sediment. *Genome Announc*. 2(6):e01184-14.

6. References

Abdallah RZ, Adel M, Ouf A, Sayed A, Ghazy MA, Alam I, *et al.* (2014). Aerobic methanotrophic communities at the Red Sea brine-seawater interface. *Front Microbiol* **5**. doi:10.3389/fmicb.2014.00487.

Antunes A, Eder W, Fareleira P, Santos H, Huber R. (2003). *Salinisphaera shabanensis* gen. nov., sp. nov., a novel, moderately halophilic bacterium from the brine-seawater interface of the Shaban Deep, Red Sea. *Extremophiles* **7**: 29–34.

Antunes A, França L, Rainey FA, Huber R, Nobre MF, Edwards KJ, *et al.* (2007). *Marinobacter salsuginis* sp. nov., isolated from the brine-seawater interface of the Shaban Deep, Red Sea. *Int J Syst Evol Microbiol* **57**: 1035–1040.

Antunes A, Ngugi DK, Stingl U. (2011). Microbiology of the Red Sea (and other) deep-sea anoxic brine lakes. *Environmental Microbiology Reports* **3**: 416–433.

Antunes A, Rainey FA, Wanner G, Taborda M, Pätzold J, Nobre MF, *et al.* (2008). A new lineage of halophilic, wall-less, contractile bacteria from a brine-filled deep of the Red Sea. **190**: 3580–3587.

Antunes A, Taborda M, Huber R, Moissl C, Nobre MF, da Costa MS. (2008). *Halorhabdus tiamatea* sp. nov., a non-pigmented, extremely halophilic archaeon from a deep-sea, hypersaline anoxic basin of the Red Sea, and emended description of the genus *Halorhabdus*. *Int J Syst Evol Microbiol* **58**: 215–220.

Backer H, Schoell M. (1972). New deeps with brines and metalliferous sediments in the Red Sea. *Nat Phys Sci* **240**: 153–158.

Balnc B, Anschutz P. (1995). New stratification in the hydrothermal brine system of the Atlantis II Deep, Red Sea. **23**: 543–546.

Barns SM, Delwiche CF, Palmer JD, Pace NR. (1996). Perspectives on archaeal diversity, thermophily and monophyly from environmental rRNA sequences. *Proceedings of the National Academy of Sciences* **93**: 9188–9193.

Bonatti E. (1985). Punctiform initiation of seafloor spreading in the Red Sea during transition from a continental to an oceanic rift. *Nature* **316**: 33–37.

Borin S, Brusetti L, Mapelli F, D'Auria G, Brusa T, Marzorati M, *et al.* (2009). Sulfur cycling and methanogenesis primarily drive microbial colonization of the highly sulfidic Urania deep hypersaline basin. *Proc Natl Acad Sci USA* **106**: 9151–9156.

Borrel G, O'Toole PW, Harris HMB, Peyret P, Brugere JF, Gribaldo S. (2013). Phylogenomic data support a seventh order of methylotrophic methanogens and provide insights into the evolution of methanogenesis. *Genome Biol Evol* **5**: 1769–1780.

Borrel G, Parisot N, Harris HMB, Peyretaillade E, Gaci N, Tottey W, *et al.* (2014). Comparative genomics highlights the unique biology of Methanomassiliicoccales, a Thermoplasmatales-

related seventh order of methanogenic archaea that encodes pyrrolysine. *BMC Genomics* **15**: 679.

Brochier-Armanet C, Boussau B, Gribaldo S, Forterre P. (2008). Mesophilic Crenarchaeota: proposal for a third archaeal phylum, the Thaumarchaeota. *Nature Reviews Microbiology* **6**: 245–252.

Brugère J-F, Borrel G, Gaci N, Tottey W, O'Toole PW, Malpuech-Brugère C. (2014). Archaebiotics: proposed therapeutic use of archaea to prevent trimethylaminuria and cardiovascular disease. *Gut Microbes* **5**: 5–10.

Bult CJ, White O, Olsen GJ, Zhou L, Fleischmann RD, Sutton GG, *et al.* (1996). Complete genome sequence of the methanogenic archaeon, *Methanococcus jannaschii*. *Science* **273**: 1058–1073.

Canfiel DE, Kristensen E, Thamdrup B. (2005). Chapter 10 The Methane Cycle. *Advances in Marine Biology Vol 48* 1–36.

Castelle CJ, Wrighton KC, Thomas BC, Hug LA, Brown CT, Wilkins MJ, *et al.* (2015). Genomic expansion of domain archaea highlights roles for organisms from new phyla in anaerobic carbon cycling. *Curr Biol* **25**: 690–701.

Cochran JR, Martinez F, Steckler MS, Hobart MA. (1986). Conrad deep: a new northern Red Sea deep. *Earth Planet Sci Lett* **78**: 18–32.

Conrad R. (2009). The global methane cycle: recent advances in understanding the microbial processes involved. *Environmental Microbiology Reports* **1**: 285–292.

Dridi B, Fardeau M-L, Ollivier B, Raoult D, Drancourt M. (2012). *Methanomassiliicoccus luminyensis* gen. nov., sp nov., a methanogenic archaeon isolated from human faeces. *Int J Syst Evol Microbiol* **62**: 1902–1907.

Eder W, Jahnke LL, Schmidt M, Huber R. (2001). Microbial diversity of the brine-seawater interface of the Kebrit Deep, Red Sea, studied via 16S rRNA gene sequences and cultivation methods. *Appl Environ Microbiol* **67**: 3077–3085.

Eder W, Ludwig W, Huber R. (1999). Novel 16S rRNA gene sequences retrieved from highly saline brine sediments of Kebrit deep, red Sea. *Arch Microbiol* **172**: 213–218.

Eder W, Schmidt M, Koch M, Garbe-Schönberg D, Huber R. (2002). Prokaryotic phylogenetic diversity and corresponding geochemical data of the brine-seawater interface of the Shaban Deep, Red Sea. *Environ Microbiol* **4**: 758–763.

Edwards AJ, Head SM. (1987). Red Sea - (Key Environments). International union for conservation of nature and natural resources.

Ferry JG. (2011). Fundamentals of methanogenic pathways that are key to the biomethanation of complex biomass. *Curr Opin Biotechnol* **22**: 351–357.

Ferry JG. (2010). How to make a living by exhaling methane. *Annu Rev Microbiol* **64**: 453–473.

Ferry JG. (2001). Methanogenesis Biochemistry. John Wiley & Sons, Ltd: Chichester, UK doi:10.1038/npg.els.0000573.

Ferry JG. (1994). Methanogenesis: ecology, physiology, biochemistry and genetics. Chapman&Hall.

Fiala G, Woese CR, Langworthy TA, Stetter KO. (1990). Flexistipes sinusarabici, a novel genus and species of eubacteria occurring in the Atlantis II Deep brines of the Red Sea. *Arch Microbiol* **154**: 120–126.

Finster K, Tanimoto Y, Bak F. (1992). Fermentation of methanethiol and dimethylsulfide by a newly isolated methanogenic bacterium. *Arch Microbiol* **157**: 425–430.

Garcia JL, Patel BK, Ollivier B. (2000). Taxonomic, phylogenetic, and ecological diversity of

methanogenic Archaea. *Anaerobe* **6**: 205–226.

Garrity GM, Bell JA, Lilburn TG. (2004). Taxonomic outline of the Procaryotes. Bergey's Manual of Systematic Bacteriology.

Gaston MA, Jiang R, Krzycki JA. (2011). Functional context, biosynthesis, and genetic encoding of pyrrolysine. *Current Opinion in Microbiology* **14**: 342–349.

Guan Y, Hikmawan T, Antunes A, Ngugi D, Stingl U. (2015). Diversity of methanogens and sulfate-reducing bacteria in the interfaces of five deep-sea anoxic brines of the Red Sea. *Research in Microbiology*. doi:10.1016/j.resmic.2015.07.002.

Halboth S, Klein A. (1992). Methanococcus voltae harbors four gene clusters potentially encoding two [NiFe] and two [NiFeSe] hydrogenases, each of the cofactor F420-reducing or F420-non-reducing types. *Molec Gen Genet* **233**: 217–224.

Hao B, Gong WM, Ferguson TK, James CM, Krzycki JA, Chan MK. (2002). A new UAG-encoded residue in the structure of a methanogen methyltransferase. *Science* **296**: 1462–1466.

Hedderich R, Whitman WB. (2006). Physiology and biochemistry of the methane-producing archaea. In: *The Prokaryotes*, Springer New York: New York, NY, pp 1050–1079.

Heinemann IU, O'Donoghue P, Madinger C, Benner J, Randau L, Noren CJ, *et al.* (2009). The appearance of pyrrolysine in tRNAHis guanylyltransferase by neutral evolution. *Proceedings of the National Academy of Sciences* **106**: 21103–21108.

Huber H, Hohn MJ, Rachel R, Fuchs T, Wimmer VC, Stetter KO. (2002). A new phylum of Archaea represented by a nanosized hyperthermophilic symbiont. *Nature* **417**: 63–67.

Jones JB, Dilworth GL, Stadtman TC. (1979). Occurrence of selenocysteine in the selenium-dependent formate dehydrogenase of *Methanococcus vannielii. Arch Biochem Biophys* **195**: 255–260.

Jones JB, Stadtman TC. (1981). Selenium-dependent and selenium-independent formate dehydrogenases of Methanococcus vannielii. Separation of the two forms and characterization of the purified selenium-independent form. *J Biol Chem* **256**: 656–663.

Kozubal MA, Romine M, Jennings RD, Jay ZJ, Tringe SG, Rusch DB, *et al.* (2013). Geoarchaeota: a new candidate phylum in the Archaea from high-temperature acidic iron mats in Yellowstone National Park. *ISME J* **7**: 622–634.

Krzycki JA. (2004). Function of genetically encoded pyrrolysine in corrinoid-dependent methylamine methyltransferases. *Curr Opin Chem Biol* **8**: 484–491.

La Cono V, Smedile F, Bortoluzzi G, Arcadi E, Maimone G, Messina E, *et al.* (2011). Unveiling microbial life in new deep-sea hypersaline Lake Thetis. Part I: Prokaryotes and environmental settings. *Environ Microbiol* **13**: 2250–2268.

Lazar CS, Parkes RJ, Cragg BA, L'Haridon S, Toffin L. (2011). Methanogenic diversity and activity in hypersaline sediments of the centre of the Napoli mud volcano, Eastern Mediterranean Sea. *Environ Microbiol* **13**: 2078–2091.

Liu Y, Beer LL, Whitman WB. (2012). Methanogens: a window into ancient sulfur metabolism. *Trends in Microbiology* **20**: 251–258.

Liu Y, Whitman WB. (2008). Metabolic, phylogenetic, and ecological diversity of the methanogenic archaea. *Annals of the New York Academy of Sciences* **1125**: 171–189.

Lopez-Garcia P, Zivanovic Y, Deschamps P, Moreira D. (2015). Bacterial gene import and mesophilic adaptation in archaea. *Nature Reviews Microbiology* **13**: 447–456.

Lowe DC. (2006). Global change: a green source of surprise. *Nature* **439**: 148–149.

Madigan MT, Martinko JM, Dunlap PV, Clark DP. (2008). Brock Biology of microorganisms 12th edn. Microbiology.

Mathrani IM, Boone DR, Mah RA, Fox GE, Lau PP. (1988). *Methanohalophilus zhilinae* sp.

nov., an alkaliphilic, halophilic, methylotrophic methanogen. *Int J Syst Bacteriol* **38**: 139–142.

McGenity TJ. (2010). Methanogens and methanogenesis in hypersaline environments. Springer Berlin Heidelberg: Berlin, Heidelberg.

Meng J, Xu J, Qin D, He Y, Xiao X, Wang F. (2014). Genetic and functional properties of uncultivated MCG archaea assessed by metagenome and gene expression analyses. *ISME J* **8**: 650–659.

Meurant G. (1976). Tectonics and Metamorphism.

Mori K, Iino T, Suzuki K-I, Yamaguchi K, Kamagata Y. (2012). Aceticlastic and NaCl-Requiring Methanogen 'Methanosaeta pelagica' sp. nov., Isolated from Marine Tidal Flat Sediment. *Appl Environ Microbiol* **78**: 3416–3423.

Narasingarao P, Podell S, Ugalde JA, Brochier-Armanet C, Emerson JB, Brocks JJ, *et al.* (2012). De novo metagenomic assembly reveals abundant novel major lineage of Archaea in hypersaline microbial communities. *ISME J* **6**: 81–93.

Ngugi DK, Blom J, Alam I, Rashid M, Ba-Alawi W, Zhang G, *et al.* (2014). Comparative genomics reveals adaptations of a halotolerant thaumarchaeon in the interfaces of brine pools in the Red Sea. *ISME J*. doi:10.1038/ismej.2014.137.

Ollivier B, Fardeau ML, Cayol JL, Magot M, Patel BK, Prensier G, *et al.* (1998). *Methanocalculus halotolerans* gen. nov., sp. nov., isolated from an oil-producing well. *Int J Syst Bacteriol* **48 Pt 3**: 821–828.

Oren A. (1999). Bioenergetic aspects of halophilism. *Microbiol Mol Biol Rev* **63**: 334–348.

Oren A. (2002). Diversity of halophilic microorganisms: Environments, phylogeny, physiology, and applications. *J Ind Microbiol Biotech* **28**: 56–63.

Oren A. (2011). Thermodynamic limits to microbial life at high salt concentrations. *Environ Microbiol* **13**: 1908–1923.

Paul K, Nonoh JO, Mikulski L, Brune A. (2012). 'Methanoplasmatales': Thermoplasmatales-related archaea in termite guts and other environments are the seventh order of methanogens.

Prat L, Heinemann IU, Aerni HR, Rinehart J, O'Donoghue P, Söll D. (2012). Carbon source-dependent expansion of the genetic code in bacteria. *Proceedings of the National Academy of Sciences* **109**: 21070–21075.

Rinke C, Schwientek P, Sczyrba A, Ivanova NN, Anderson IJ, Cheng J-F, *et al.* (2013). Insights into the phylogeny and coding potential of microbial dark matter. *Nature* **499**: 431–437.

Rosenberg E, DeLong EF, Lory S, Stackebrandt E. (2014). The prokaryotes : other major lineages of Bacteria and the Archaea.

Rother M, Krzycki JA. (2010). Selenocysteine, Pyrrolysine, and the Unique Energy Metabolism of Methanogenic Archaea. *Archaea* **2010**: –14.

Sakai S, Imachi H, Hanada S, Ohashi A, Harada H, Kamagata Y. (2008). *Methanocella paludicola* gen. nov., sp. nov., a methane-producing archaeon, the first isolate of the lineage 'Rice Cluster I', and proposal of the new archaeal order Methanocellales ord. nov. *Int J Syst Evol Microbiol* **58**: 929–936.

Schmidt M, Al-Farawati R, Botz R. (2015). Geochemical classification of brine-filled Red Sea deeps. In: *The Red Sea*, Springer Earth System Sciences, Springer Berlin Heidelberg: Berlin, Heidelberg, pp 219–233.

Siam R, Mustafa GA, Sharaf H, Moustafa A, Ramadan AR, Antunes A, *et al.* (2012). Unique prokaryotic consortia in geochemically distinct sediments from Red Sea Atlantis II and Discovery Deep brine pools Mormile, MR (ed). *PLoS ONE* **7**. doi:10.1371/journal.pone.0042872.

Sorgenfrei O, Linder D, Karas M, Klein A. (1993). A novel very small subunit of a selenium

containing [NiFe] hydrogenease of *Methanococcus voltae* is postranslationally processed by cleavage at a defined position. *European Journal of Biochemistry* **213**: 1355–1358.

Spang A, Saw JH, Jørgensen SL, Zaremba-Niedzwiedzka K, Martijn J, Lind AE, *et al.* (2015). Complex archaea that bridge the gap between prokaryotes and eukaryotes. *Nature*. doi:10.1038/nature14447.

Spring S, Scheuner C, Lapidus A, Lucas S, Glavina Del Rio T, Tice H, *et al.* (2010). The genome sequence of *Methanohalophilus mahii* SLP(T) reveals differences in the energy metabolism among members of the *Methanosarcinaceae* inhabiting freshwater and saline environments. *Archaea* **2010**: 690737.

Stock T, Selzer M, Rother M. (2010). In vivo requirement of selenophosphate for selenoprotein synthesis in archaea. *Molecular Microbiology* **75**: 149–160.

Swift SA, Bower AS, Schmitt RW. (2012). Vertical, horizontal, and temporal changes in temperature in the Atlantis II and Discovery hot brine pools, Red Sea. *Deep Sea Research Part I:* ….

Thauer RK, Kaster A-K, Seedorf H, Buckel W, Hedderich R. (2008). Methanogenic archaea: ecologically relevant differences in energy conservation. *Nature Reviews Microbiology* **6**: 579–591.

Ueno Y, Yamada K, Yoshida N, Maruyama S, Isozaki Y. (2006). Evidence from fluid inclusions for microbial methanogenesis in the early Archaean era. *Nature* **440**: 516–519.

van der Wielen PWJJ. (2005). The enigma of prokaryotic life in deep hypersaline anoxic basins. *Science* **307**: 121–123.

Vorholt JA, Vaupel M, Thauer RK. (1997). A selenium-dependent and a selenium-independent formylmethanofuran dehydrogenase and their transcriptional regulation in the hyperthermophilic *Methanopyrus kandleri*. *Molecular Microbiology* **23**: 1033–1042.

Wang Y, Cao H, Zhang G, Bougouffa S, Lee OO, Al-Suwailem A, *et al.* (2013). Autotrophic microbe metagenomes and metabolic pathways differentiate adjacent red sea brine pools. *Sci Rep* **3**: 1748.

Watkins AJ, Roussel EG, Parkes RJ, Sass H. (2014). Glycine betaine as a direct substrate for methanogens (*Methanococcoides* spp.). *Appl Environ Microbiol* **80**: 289–293.

Watkins AJ, Roussel EG, Webster G, Parkes RJ, Sass H. (2012). Choline and N,N-dimethylethanolamine as direct substrates for methanogens. *Appl Environ Microbiol* **78**: 8298–8303.

Watson SW, Waterbury JB. (1969). The sterile hot brines of the Red Sea. In: *Hot Brines and Recent Heavy Metal Deposits in the Red Sea*, Springer Berlin Heidelberg: Berlin, Heidelberg, pp 272–281.

Whitman WB, Bowen TL, Boone DR. (2006). The methanogenic bacteria. 165–207.

Woese CR, Fox GE. (1977). Phylogenetic structure of the prokaryotic domain: the primary kingdoms. *Proceedings of the National Academy of Sciences* **74**: 5088–5090.

Woese CR, Kandler O, Wheelis ML. (1990). Towards a natural system of organisms: proposal for the domains Archaea, Bacteria, and Eucarya. *Proc Natl Acad Sci USA* **87**: 4576–4579.

Yakimov MM, La Cono V, La Spada G, Bortoluzzi G, Messina E, Smedile F, *et al.* (2014). Microbial community of the deep-sea brine Lake Kryos seawater-brine interface is active below the chaotropicity limit of life as revealed by recovery of mRNA. *Environ Microbiol* **17**: n/a–n/a.

Yakimov MM, La Cono V, Slepak VZ, La Spada G, Arcadi E, Messina E, *et al.* (2013). Microbial life in the Lake Medee, the largest deep-sea salt-saturated formation. **3**. doi:10.1038/srep03554.

Zhilina TN. (1986). Methanogenic bacteria from hypersaline environments. *Syst Appl Microbiol*

7: 216–222.

Zhilina TN, Zavarzin GA. (1987). *Methanohalobium evestigatum*, n. gen., n. sp. the extremely halophilic methanogenic archaebacterium. *Doklady Akademii nauk SSSR* **293**: 464–468.

1 Chapter I DIVERSITY OF METHANOGENS AND SULFATE-REDUCING BACTERIA IN THE INTERFACES OF FIVE DEEP-SEA ANOXIC BRINES OF THE RED SEA

This chapter was published as:

Guan Y*, Hikmawan T*, Antunes A, Ngugi DK, Stingl U. 2015.Diversity of methanogens and sulfate-reducing bacteria in the interfaces of five deep-sea anoxic brines of the Red Sea. *Research in Microbiology*. http://dx.doi.org/10.1016/j.resmic.2015.07.002.

Abstract

Oceanic deep hypersaline anoxic basins (DHABs) are characterized by drastic changes in physico-chemical conditions in the transition from overlaying seawater to brine body. Brine-seawater interfaces (BSIs) of several DHABs across the Mediterranean Sea have been shown to possess methanogenic and sulfate-reducing activities, yet no systematic studies have been conducted to address the potential functional diversity of methanogenic and sulfate-reducing communities in the Red Sea DHABs. Here, we evaluated the relative abundance of Bacteria and Archaea using quantitative PCR and conducted phylogenetic analyses of nearly full-length 16S rRNA genes as well as functional marker genes encoding the alpha subunits of methyl-coenzyme M reductase (*mcrA*) and dissimilatory sulfite reductase (*dsrA*). Bacteria predominated over Archaea in most locations, the majority of which being affiliated with *Deltaproteobacteria*, while *Thaumarchaeota* were the most prevalent Archaea in all sampled locations. The upper convective layers of Atlantis II Deep, which bear increasingly harsh environmental conditions, were dominated by members of the class *Thermoplasmata* (Marine Benthic Group E and Mediterranean Sea Brine Lakes Group 1). Our study revealed unique microbial compositions, the presence of niche-specific groups, and collectively a higher diversity of the sulfate-reducing communities compared to the methanogenic communities in all five studied locations.

1.1 Introduction

Hypersaline water bodies at the bottom of the ocean (brine pools) are present in the Mediterranean Sea, the Gulf of Mexico, and the Red Sea (Antunes *et al.*, 2011). In the Red Sea, a total of 25 such deep-sea hypersaline brine pools have been discovered at depths ranging from 1,196 to 2,850 meters below sea level (Antunes *et al.*, 2011; Backer and Schoell, 1972). These environments are extremely saline (up to 26% salinity), anoxic, rich in heavy metals, and characterized by drastic changes in physicochemical conditions when compared to the overlaying seawater (Hartmann *et al.*, 1998).

The interface between the brine pools and the seawater (BSI) represents a highly peculiar environment that harbors a high microbial diversity and biomass (Borin *et al.*, 2009; Bougouffa *et al.*, 2013; Eder *et al.*, 2002). The increase in microbial biomass can be explained by the drastic changes in density, which acts as an in situ particle trap for debris sinking through the water column, thus increasing the concentrations of available nutrients (Eder *et al.*, 2002; LaRock *et al.*, 1979). In addition, the BSI is also characterized by sharp changes in physicochemical parameters including salinity, oxygen concentration, temperature, and redox potential, all of which provide a large variety of environmental niches for different metabolic groups (Daffonchio *et al.*, 2006; Sass *et al.*, 2001). The microbiology of the BSIs of some of the Red Sea brine pools has been explored with a combination of cultivation-dependent (Antunes *et al.*, 2003; 2007) and molecular–based methods (Eder *et al.*, 2001; Abdallah *et al.*, 2014). Previous studies based on 16S rRNA gene

sequences uncovered novel groups of Archaea and Bacteria inhabiting the BSI of Shaban Deep and Kebrit Deep of the Red Sea (Eder *et al.*, 2002; 2001).

Microbial community studies in the Mediterranean DHABs revealed that diverse biogeochemical processes apparently co-occur in the BSI (Daffonchio *et al.*, 2006). Other investigations reported on the importance of methanogenesis and sulfur cycling in these environments (Borin *et al.*, 2009; van der Wielen and Heijs, 2007). These findings were also corroborated by recent metagenomic studies, where pathways for methanogenesis and/or sulfate reduction were detected in brines from the Red Sea and in DHABs in the Mediterranean Sea (Wang *et al.*, 2011; 2013). Additionally, unique microbial communities were found to thrive in the sediments of two brine pools in the Red Sea, and many of the reported microorganisms are hypothesized to play a dominant role in the methane and sulfur cycle, based on their phylogenetic affiliations (Siam *et al.*, 2012).

Taken together, methanogenesis and sulfate reduction could thus be considered very important biogeochemical processes in deep-sea brines (Borin *et al.*, 2009) (van der Wielen, 2005). However, the compositions of the microbial communities involved in both processes are largely unknown for the Red Sea brine pools. Considering the extreme conditions of these environments and the unique combination of physicochemical features in each individual brine pool (Antunes *et al.*, 2011), we postulated the existence of novel, niche-adapted groups of methanogens and sulfate reducers in the BSIs. Moreover, despite the micro-oxic conditions present in the BSI (Ngugi *et al.*,

2014), members of both groups are capable of tolerating minute amounts of oxygen (Angel *et al.*, 2011; Brune *et al.*, 2000; Marschall *et al.*, 1993) and could thus play an important role in these environments. Previous 454 amplicon data (Ngugi *et al.*, 2014) uncovered interesting microbial communities in the sampled sites, but as many of the relatively short sequences stem from poorly characterized groups, we decided that nearly full-length 16S rRNA gene sequences would be important to provide better phylogenetic detail and resolution on members of these groups. Therefore, we analyzed the microbial communities in the BSIs of geochemically distinct brine pools of the Red Sea, using the canonical 16S rRNA gene, as well as functional marker genes encoding for the alpha subunits of methyl-coenzyme M reductase (*mcrA*) and dissimilatory sulfite reductase (*dsrA*) to uncover the main methanogenic and sulfate-reducing communities.

1.2 Materials and methods

1.2.1 Sample collection

Water samples from the brine-seawater interfaces and the upper convective layers of the deep-sea brines were collected from the R/V *Aegaeo* during the 3rd KAUST Red Sea Expedition in November 2011 using a rosette sampler equipped with 10-l Niskin bottles and a CTD unit for monitoring salinity, temperature, transmission, oxygen, and pressure (Idronaut, Italy). Large volumes (ca. 200 l) of sample were collected from Atlantis II Deep BSI (Ai); first, second, and third upper-convective layer of Atlantis II Deep (labeled as A-UCL1, A-UCL2, A-UCL3, respectively), Discovery Deep BSI (Di), Erba Deep BSI (Ei), Kebrit Deep BSI (Ki), and Nereus Deep BSI (Ni) (Table 1). During

sampling, we have avoided mixing between the seawater and the brine samples by carefully controlling the depth of the CTD and sampler when triggering the closure of each Niskin bottle to ensure sampling of desired layers. Furthermore, prior to sample collection on deck, we measured the salinities at the top and bottom of each individual Niskin bottle using a handheld refractometer (MASTER Refractometer, Atago, Japan) to confirm that the salinities of the samples matched the expected values of the targeted layers. Samples were then concentrated using a Tangential Flow Filtration (TFF) as described elsewhere (Ngugi *et al.*, 2014). Methane and carbon dioxide concentrations in the samples were determined on a gas chromatograph equipped with FID, and a RTX1-60m capillary column (ø=0.52mm), using N_2 as carrier gas, by the commercial service provided by GEOMAR Helmholtz Centre for Ocean Research (Kiel, Germany).

1.2.2 DNA extraction, amplification, and sequencing of 16S rRNA genes

Nucleic acids were extracted as previously described (Ngugi and Stingl, 2012) and the concentrations of the DNA were measured in a NanoDrop (Thermo Scientific, USA). Partial 16S rRNA genes were amplified by PCR by using combinations of the archaeal-specific primer 4F (5′-TCCGGTTGATCCTGCCRG-3′) (Kendall *et al.*, 2007), or the bacteria-specific primer 27F (5′-AGAGTTTGATCMTGGCTCAG-3′) paired with the universal primer 1492R (5′-GGTTACCTTGTTACGACTT-3′) (Lane, 1991). The primers were chosen to produce sequences with maximum length. *In silico* testing using Silva-TestPrime (http://www.arb-silva.de/search/testprime/ (Klindworth *et al.*, 2013) with one allowed mismatch indicated a coverage for Bacteria of 65.5%

and a coverage for Archaea of 46.6%, while three allowed mismatches indicated a coverage for Bacteria of 71% and a coverage for Archaea of 53.1%. In addition, the above primers have a good coverage of major taxa reported in a previous study using 454 amplicon data (Ngugi *et al.*, 2014). The PCR conditions for archaeal 16S rRNA genes were: an initial denaturation of 5 min at 94 °C, 30 cycles of 1 min at 94 °C, 1 min at 55 °C, and 1.5 min at 72 °C, and then a final extension step of 7 min at 72 °C. The conditions for bacterial PCR were 3 min at 94 °C, 35 cycles of 1 min at 94 °C, 1 min at 53 °C, and 1.5 min at 72 °C, and then 7 min at 72°C. Purified PCR products were cloned into PCR®2.1 TOPO vectors (Invitrogen) according to the manufacturer's instructions. All clones with inserts from each library (856 and 1040 for archaeal and bacterial libraries, respectively) were selected for plasmid extraction and bi-directional sequencing on an ABI 3730 × l Capillary Sequencer at the Biosciences Core Laboratory at KAUST. Raw 16S rRNA gene sequences were quality checked, trimmed and assembled using Sequencher v.4.9 (Gene Codes Corporation).

1.2.3 Diversity and phylogenetic analysis

Assembled archaeal and bacterial 16S rRNA sequences were aligned and analyzed using mothur v.1.31, yielding operational taxonomic units (OTUs) grouped at 97% sequence identity level (Schloss *et al.*, 2009). Potential chimeric sequences were removed using the UCHIME 4.2 package (Edgar *et al.*, 2011), and diversity indeces (Shannon) and estimated sample coverage (Good's coverage) were calculated as implemented in mothur. Sequence alignments of the resulting representative OTUs (62 for Archaea and 281 for

Bacteria) and closely related sequences recovered from GenBank using BLASTn (Altschul *et al.*, 1997), were done automatically using the SINA aligner (http://www.arb-silva.de/aligner/) against the SILVA SSU 115 database (Pruesse *et al.*, 2007). The SILVA-aligned sequences were then used to construct phylogenetic trees with a maximum likelihood algorithm using bootstrap analysis (1000 samples) to validate support for clades as implemented in ARB v.5.3 (Ludwig *et al.*, 2004).

1.2.4 Functional gene analysis

Partial *mcrA* and *dsrA* genes were amplified based on previously described methods (Springer *et al.*, 1995; Pester *et al.*, 2010). Purification of PCR products, clone library construction and sequencing protocols were the same as the ones used for 16S rRNA genes described above. The OTUs of *mcrA* and *dsrA* genes were generated at a 6% distance cutoff using FunGene (http://fungene.cme.msu.edu/) (Fish *et al.*, 2013). Deduced amino acid sequences of both genes were aligned with ClustalW (Thompson *et al.*, 1994). The alignments of *mcrA* and *dsrA* genes were used for phylogenetic analyses with a maximum-likelihood algorithm (amino acids substitution model: LG for *mcrA*, and LG+G for *dsrA* genes) and 1000 bootstraps as implemented in Geneious Pro version 7.1 (Biomatters Ltd.) and MEGA version 6.06 (Tamura *et al.*, 2011), respectively.

1.2.5 Quantification of gene copy numbers by real-time PCR

Copy numbers of total bacterial and archaeal 16S rRNA genes from each sample location were determined by quantitative real-time PCR (qPCR) using EXPRESS qPCR SuperMix (Invitrogen) and a two-step qPCR cycling program

on an ABI 7900HT Fast Real-Time PCR System instrument (Applied Biosystem). The primers Bac518F and Bac786R for Bacteria and A519F and A727R for Archaea were used as described by Park et al. (Park *et al.*, 2010). Standards were made from plasmids containing inserts of archaeal or bacterial 16S rRNA gene sequences. The efficiency for archaeal 16S rRNA primers was 99.8% and that for bacterial 16S rRNA primers was 99.5%, as estimated based on the slope of the standard curve. To allow for better comparisons among the different samples, copy numbers of genomic DNA were normalized based on ng of genomic DNA.

1.2.6 Nucleotide sequence accession numbers

The 16S rRNA gene sequences from this study were deposited in GenBank under accession numbers KJ881441–KJ882283 (Archaea), and KM018335–KM019141, KP083299– KP083370 (Bacteria), while *mcrA* and *dsrA* gene sequences were deposited under accession numbers KJ880100–KJ880274 and KM241874–KM242055, respectively.

1.3 Results and discussion

1.3.1 General microbial community structure

The ratios of the copy numbers of bacterial to archaeal 16S rRNA genes as estimated using qPCR were used as a proxy for their abundances in each sample. These ratios ranged in our samples from 0.15 to 179.12 (Table 5), and Bacteria were more abundant than Archaea in six out of eight samples. The predominance of Bacteria over Archaea thus seems to be a general trend in the BSI (and brine bodies) of brine pools from the Red Sea ((Bougouffa *et al.*, 2013; Wang *et al.*, 2013); this study) and the Mediterranean Sea (Daffonchio

et al., 2006; van der Wielen, 2005; Yakimov *et al.*, 2014; Yakimov, La Cono, *et al.*, 2007); the exceptions being the BSIs of Kebrit Deep and Atlantis II Deep (this study) and the brine layer of the Urania DHAB in the Mediterranean Sea (van der Wielen, 2005). Physico-chemical differences that might be either the reason for, or a result of, the high archaeal abundances are 1) they are highly sulfidic (in the case of Kebrit Deep and Urania DHAB; ~150 µM and 10 mM H_2S, respectively (Sass *et al.*, 2001)), and 2) the BSIs of Kebrit and Atlantis II Deep have dissolved oxygen concentrations (DO_2) that are 4–9 times higher than those at the other locations, where DO_2 is close to depletion (Table 5).

Table 5. Physical and geochemical parameters of sampling locations, bacteria/archaea ratios, and an overview of the clone libraries.

Sampling site[a]		Kebrit	Erba	Nereus	Discovery	Atlantis II			
		Ki	Ei	Ni	Di	Ai	A-UCL1	A-UCL2	A-UCL3
Latitude (N)[b]		24° 43.41'	20° 43.80'	23° 11.53'	21° 16.98'	21° 20.76'	21° 20.76'	21° 20.76'	21° 20.76'
Longitude (E)[b]		36° 16.63'	38° 10.98'	37° 25.09'	38° 3.18'	38° 4.68'	38° 4.68'	38° 4.68'	38° 4.68'
Depth (m)[b]		1467	2381	2432	2038	1998	2006	2025	2048
Salinity (%)[b]		18.2	9.8	15.4	13.8	5.6	8.6	11.2	15.4
H_2S (µmol/l)[b]		149.8	b.d.l.	b.d.l.	b.d.l.	b.d.l.	b.d.l.	b.d.l.	b.d.l.
CH_4 (ppmV)		28955.5	15.5	23	53	b.d.l.	85.5	330	977
CO_2 (ppmV)		103731	<400	<400	791.5	7878	18953	19288	12226
16S rRNA gene	*Bacteria*	6.6	20.6	11.41	457.69	48.05	59.11	1.44	1.27
copy nr. ($x10^4$)	*Archaea*	42.69	1.03	0.08	54.81	77.8	0.33	0.39	0.09
Bacteria/Archaea 16S rRNA gene copy nr. ratio[c]		0.15	20	142.63	8.35	0.62	179.12	3.69	14.11
Bacteria/Archaea ratio after normalization[d]		0.07	9.07	64.07	3.79	0.28	81.26	1.67	6.40
Nr. of 16S rRNA gene clones [nr. of OTUs]	*Bacteria*	250 [73]	189 [91]	104 [53]	121 [53]	92 [18]	181 [59]	-[e]	23 [12]
	Archaea	160 [22]	155 [18]	84 [9]	118 [9]	95 [7]	96 [6]	100 [12]	35 [12]
Shannon index	*Bacteria*	3.03	4.03	3.65	3.3	1.84	3.43	-	2.24
	Archaea	1.44	1.45	1.1	1.11	0.64	0.87	1.86	2.36
Good's coverage	*Bacteria*	0.71	0.68	0.68	0.70	0.85	0.80	-	0.48
	Archaea	0.93	0.97	0.96	1	0.97	1	0.99	0.94
Nr. of clones of functional genes [OTUs]	*mcrA*	12 [3]	152 [2]	-	-	-	-	20 [1]	15 [1]

	dsrA	42 [8]	112 [13]	31 [7]	18 [3]	5 [3]	12 [5]	-	-
Library coverage	*mcrA*	0.75	0.99	-	-	-	-	0.95	0.93
	dsrA	0.81	0.88	0.77	0.83	0.4	0.58	-	-

[a] Abbreviations for sampling sites: Ki, Kebrit Deep BSI; Ei, Erba Deep BSI; Ni, Nereus Deep BSI; Di, Discovery Deep BSI; Ai, Atlantis II Deep BSI; A-UCL1, A-UCL2, and A-UCL3, the first, second, and third upper-convective layer, respectively.

[b] For physicochemical data please refer to Ngugi et al. 2014

[c] To allow for a better comparison among the different samples, copy numbers of genomic DNA were normalized based on ng of genomic DNA.

[d] Archaeal and bacterial abundance ratios were estimated based on the qPCR results and the average 16S rRNA gene copy 4.1 per cell in Bacteria and 1.86 per cell in Achaea (Lee et al. 2009).

[e] The bacterial library of A-UCL2 was not generated because of technical issues.

Phylogenetic analysis based on 16S rRNA gene sequences revealed a high microbial diversity in the brine interfaces of geochemically distinct brine pools of the Red Sea (Figure 6). A total of 843 archaeal (>1000 bp length) and 960 bacterial (>1400 bp length) non-chimeric 16S rRNA gene sequences were obtained from the BSIs of the five brines and from the subsequent three upper convective layers of Atlantis II Deep (A-UCL1, A-UCL2, and A-UCL3). These sequences were clustered into 62 and 281 OTUs (at 97% sequences identity level) for archaeal and bacterial genes, respectively. Archaeal sequences were primarily affiliated with the phyla *Thaumarchaeota* (60%) and *Euryarchaeota* (37%), while the bacterial sequences encompassed diverse lineages. Our findings are in general consistent with previous studies (Bougouffa *et al.*, 2013; Ngugi *et al.*, 2014).

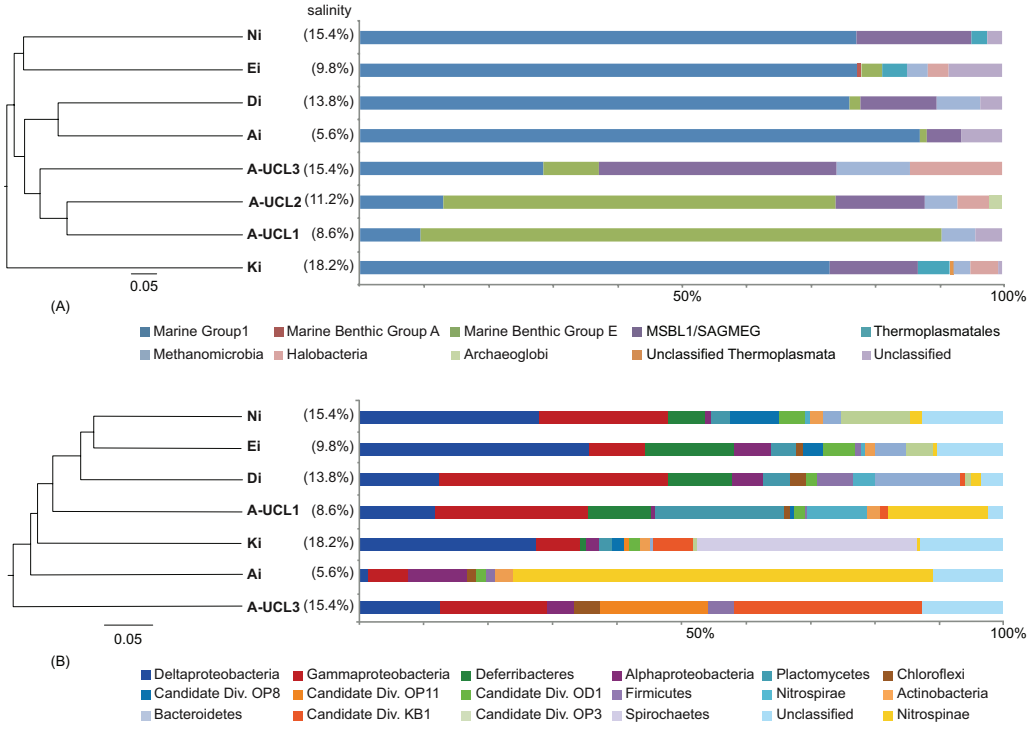


Figure 6. Taxonomic classification and relative abundance of archaeal (A) and bacterial (B) communities in the brine-seawater interfaces of five different brine pools of the Red Sea. A total of 843 archaeal and 960 bacterial 16S rRNA gene fragments were classified using Mothur based on SILVA database at 97% cutoff. The cluster dendrogram illustrates the linked hierarchical clustering of different environments based on the relative abundance of the OTUs in each sampling location. (**Ai**, Atlantis II Deep BSI; A-UCL1, A-UCL2, and A-UCL3, the first, second, and third upper convective layer, respectively; **Di**, Discovery Deep BSI; **Ei**, Erba Deep BSI; **Ki**, Kebrit Deep BSI; **Ni**, Nereus Deep BSI. MSBL1, Mediterranean Sea Brine Lakes Group 1; SAGMEG, South African Goldmine Euryarchaeotal Group).

Hierarchical cluster analysis of the archaeal and bacterial 16S rRNA gene based on the Jaccard similarity index indicated similarities and differences among microbial communities of each location (Figure 6). The archaeal community in Kebrit Deep was clearly distinctive from those in the other locations. Archaeal communities in the three upper convective layers of Atlantis II Deep, BSIs of Nereus Deep and Erba Deep, and BSIs of Atlantis II Deep and Discovery Deep formed three separate clusters. In contrast with the results of the archaeal communities, we found no apparent clustering of the

bacterial communities. Though, a highly stratified bacterial community profile was observed in the multi-layered Atlantis II Deep, concurrent with previous reports (Bougouffa *et al.*, 2013).

This is the first detailed phylogenetic study of nearly full-length prokaryotic 16S rRNA gene sequences from BSIs of Erba and Nereus Deep. Surprisingly, despite their salinity differences (9.8% vs. 15.4 %), they both harbor very similar microbial communities (Figure 6). The primary archaeal taxon was Marine Group I *Thaumarchaeota* (77%), while members of the class *Deltaproteobacteria* dominated their bacterial clone libraries (35.5% in Erba, and 27.9% in Nereus). In general, the microbial community compositions based on clone libraries were consistent with previous findings using an amplicon sequencing approach (Ngugi *et al.*, 2014). Additionally, six bacterial OTUs accounting for ~8% of all sequences were present in both brine pools, and are affiliated mostly with sulfate-reducing or sulfur-oxidizing taxa (e.g., *Deltaproteobacteria* and SAR324; Table S1).

1.3.2 Detailed analysis of archaeal communities

Both *Thaumarchaeota* and *Thermoplasmata* were ubiquitous in all five brine-seawater interfaces (BSI) and the three upper convective layers of Atlantis II Deep (A-UCL1, A-UCL2, and A-UCL3). This suggests that they are important components of deep-sea brine environments, and presumably possess adaptations to thrive in these gradient environments. *Thaumarchaeota* was the predominant group in archaeal clone libraries of all the five BSI samples (73–87%). The archaeal community composition in the Atlantis II BSI was different from the communities found in the subsequent convective layers. Clone

libraries showed that the most abundant members in those convective layers belonged to the class *Thermoplasmata* (46-81%). Class *Methanomicrobia*-related sequences were also present (six out of eight sampled locations), but constituted only a small proportion of the archaeal communities (2.5%–11.4%; Figure 6A). Additionally, a variety of archaeal lineages were found in the investigated Red Sea brine pool BSIs, albeit at low abundances, such as *Archaeoglobi*, *Halobacteria*, Marine Benthic Group A and D, Marine Group III, MSP41, CCA47 cluster and VC2.1 Arc6, Terrestrial Miscellaneous Group, SM1K20 group, South African Goldmine Euryarchaeotal Group (SAGMEG), Deep-Sea Euryarchaeotic Group, Miscellaneous Euryarchaeotic Group (MEG) and Terrestrial Hot Spring Crenarchaeotic Group (THSCG) (Figure S2). Many of the phylogenetic lineages of Archaea retrieved in this study are uncultured at present. A few of these lineages have been previously detected from the sediment samples of Atlantis II Deep and Discovery Deep of the Red Sea in a pyrosequencing approach using 16S rRNA genes (Siam *et al.*, 2012).

Although we cannot rule out the possibility of reporting DNA sequences of non-metabolically active cells due to long-term preservation of DNA and cells in the deep-sea hypersaline environments, our recently published results (Ngugi *et al.*, 2014) have shown that most of the thaumarchaeal sequences retrieved from the exact same sampled locations (BSIs) belonged to a Marine Group I phylotype that is absent in the overlaying water column. Still, this might not apply to all retrieved sequences. The cultivated species in the phylum *Thaumarchaeota* are autotrophic ammonia-oxidizing archaea (Konneke *et al.*, 2005), but a mixotrophic lifestyle in the dark ocean has also

been suggested (Swan *et al.*, 2014). Considering their abundance and the halotolerant genomic features of thaumarchaeal single-cells from BSIs of Red Sea brines (Ngugi *et al.*, 2014), we reconfirm previous findings that the BSI populations might play significant roles in the nitrogen and carbon cycles in the deep-sea brine interfaces with special adaptation to the hypersaline environments.

Detected members of the *Thermoplasmata* were mostly associated with the Candidate division MSBL1 (or Mediterranean Sea Brine Lakes Group 1) and MBGE (Marine Benthic Group E). Considerable proportions (5%–37%) of sequences related to MSBL1 were found in most clone libraries (six out of eight). In the Mediterranean DHABs, MSBL1 communities were commonly found in the lower interfaces and in the brines, and they have been assumed to be methanogenic due to a positive correlation between their abundances and the *in situ* methane concentrations (Yakimov *et al.*, 2013; La Cono *et al.*, 2011).

MBGE sequences were more abundant in the clone libraries of the upper convective layers of Atlantis II Deep than in other samples. This is consistent with previous reports indicating that they frequently occur at locations with higher temperatures (50 °C – 63 °C), correlating with the high G+C content of their 16S rRNA gene sequences (Kato *et al.*, 2009). They were often retrieved from deep-sea sediment, hydrothermal environments, chimney samples (Kato *et al.*, 2010), and iron-rich habitats (Takai *et al.*, 2008). Therefore, we speculate that their increased abundance in the upper convective layers of

Atlantis II is related to the increased temperature (52 – 65 °C) and iron concentration (24.5–70.5 μM). As no members of the MBGE clade and MSBL1 clade have been cultivated so far, their physiology and ecological roles in hypersaline deep-sea environments require future investigation.

1.3.3 Detailed analysis of bacterial communities

A wide diversity of Bacteria was observed in the BSIs of Atlantis II, Discovery, Kebrit, Nereus, and Erba Deep, corroborating previous reports on the microbial communities in the Red Sea brine pools (Bougouffa *et al.*, 2013; Eder *et al.*, 2002; 2001). The majority of bacterial 16S rRNA sequences were affiliated with *Proteobacteria* (relative abundance 37.5–89.4%), *Bacteroidetes* (0.8–13.2%), *Deferribacteres* (0.8–13.9%), and *Chloroflexi* (1.1–4.2%) (Figure 6B). The remaining bacterial sequences belong to uncultured bacterial groups with unknown physiology such as Candidate division KB1, MSBL 2, and ST12-K34 (Figure S3). These uncultivated groups were previously reported from hypersaline brines in the Red Sea and DHABs in the Mediterranean Sea (Eder *et al.*, 2001; 2002; Yakimov *et al.*, 2013).

Deltaproteobacteria was the predominant bacterial class in the BSIs of colder brines such as Kebrit, Erba, and Nereus Deep (Figure 6B). The high abundance of *Deltaproteobacteria* in the BSIs of these three brines is consistent with previous studies from the Mediterranean DHABs (Yakimov, La Cono, *et al.*, 2007; Borin *et al.*, 2009). This group is also one of the most prominent metabolically active microbial groups thriving in the chemocline of $MgCl_2$-rich Discovery Basin (Hallsworth *et al.*, 2007), hypersaline Lake Kryos (Yakimov *et al.*, 2014), and hydrothermal mud fluids of Urania DHAB, Mediterranean Sea

(Yakimov, Giuliano, *et al.*, 2007). These findings affirm the argument that sulfate reduction is one of the main metabolic processes occurring in the chemoclines of brine pools, which might be primarily performed by members of the *Deltaproteobacteria* (van der Wielen, 2005; Borin *et al.*, 2009).

On the contrary, *Deltaproteobacteria* were less abundant in the clone libraries of the hot brines such as Discovery and Atlantis II Deep. The bacterial clone libraries of Atlantis II Deep shifted from being dominated by *Nitrospinae*-like bacteria in the BSI to being dominated by *Gammaproteobacteria* in the convective layers underneath (Figure 6B). Similar patterns of gradually changing microbial communities were observed in a previous study based using pyrosequencing of the 16S rRNA genes in these two locations (Bougouffa *et al.*, 2013). According to the same study, the combination of high temperature and salinity was presumed to shape the communities in both brines.

The majority of the OTUs in the class *Deltaproteobacteria* fell into four orders: *Desulfobacterales*, *Desulfurellales*, *Desulfovibrionales*, and *Syntrophobacterales* (Figure 7). The remaining OTUs were assigned into four lineages with no cultured representatives, namely 10bav-F6, DTB120, *Candidatus* Entotheonella, and SAR 324. A large fraction of *Deltaproteobacteria* (around 33.3%) in A-UCL1 could not be classified to any known family, based on the ARB-SILVA SSU 115 database.

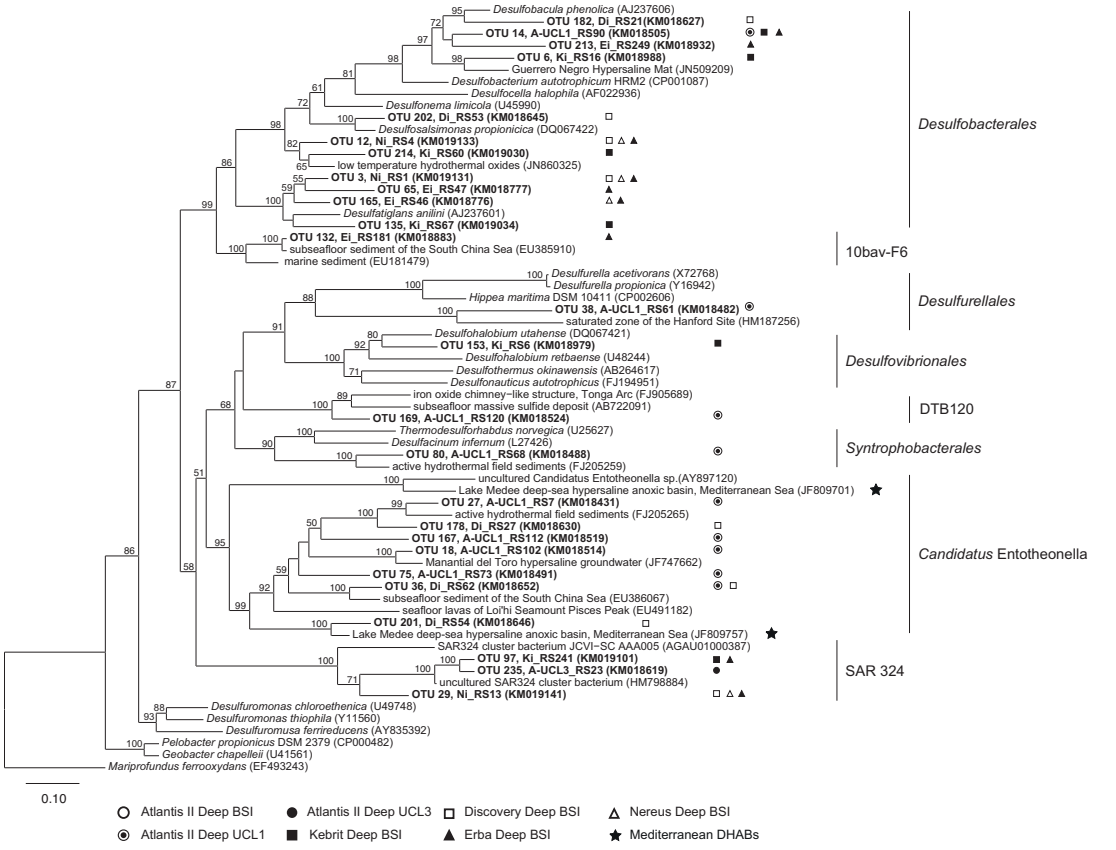


Figure 7. 16S rRNA gene-based phylogenetic tree of the *Deltaproteobacteria* group, including the representative sequences from the Atlantis II, Discovery, Erba, Kebrit, and Nereus Deep. The topology of the tree is based on maximum-likelihood algorithm with 1000 bootstraps. The scale bar represents 0.10 fixed mutation per nucleotide position. Bootstrap values above 50% are shown.

Certain bacterial groups seemed specifically adapted to high salinities. For instance, the bacterial community of the deepest convective layer of Atlantis II (A-UCL3) was dominated by Candidate division KB1, a group with no cultured representatives. This group branches between the *Aquificales* and the *Thermotogales*, and is restricted to hypersaline conditions (Eder *et al.*, 1999). Members of KB1 were initially retrieved from sediments in Kebrit Deep (Eder *et al.*, 2001) and later obtained in enrichments of BSI with high salinities obtained from Lake Medee, the Mediterranean Sea (Yakimov *et al.*, 2013). In this study, Candidate division KB1 was detected in interfaces with higher

salinity such as A-UCL3, Di, and Ki (Figure S3). The metabolic preferences of Candidate division KB1 remain unsolved although they seem to partially rely on reductive cleavage of the osmoprotectant glycine betaine, resulting in the formation of trimethylamine (TMA) and acetate (Yakimov *et al.*, 2013).

1.3.4 Molecular diversity of mcrA genes

Up to now, all described methanogenic archaea fall into the seven orders in the phylum *Euryarchaeota*: *Methanococcales*, *Methanopyrales*, *Methanobacteriales*, *Methanosarcinales*, *Methanomicrobiales*, *Methanocellales*, and the recently proposed 7th order, *Methanomassiliicoccales* (Borrel *et al.*, 2013). The reduction of CO_2, the fermentation of acetate, and the dismutation of methanol or methylamines encompass the three major methanogenic pathways (Ferry, 2001). Although methanogens play an important role in the global carbon cycling in various environments, very little is known about the methanogenic players in the Red Sea brine pools. Methyl-coenzyme M reductase is unique to methanogens and catalyzes the last step in methane formation. Genes encoding the α subunit of this enzyme (*mcrA*) have been employed as a specific marker to detect and differentiate methanogenic and anaerobic methanotrophic communities (Friedrich, 2005).

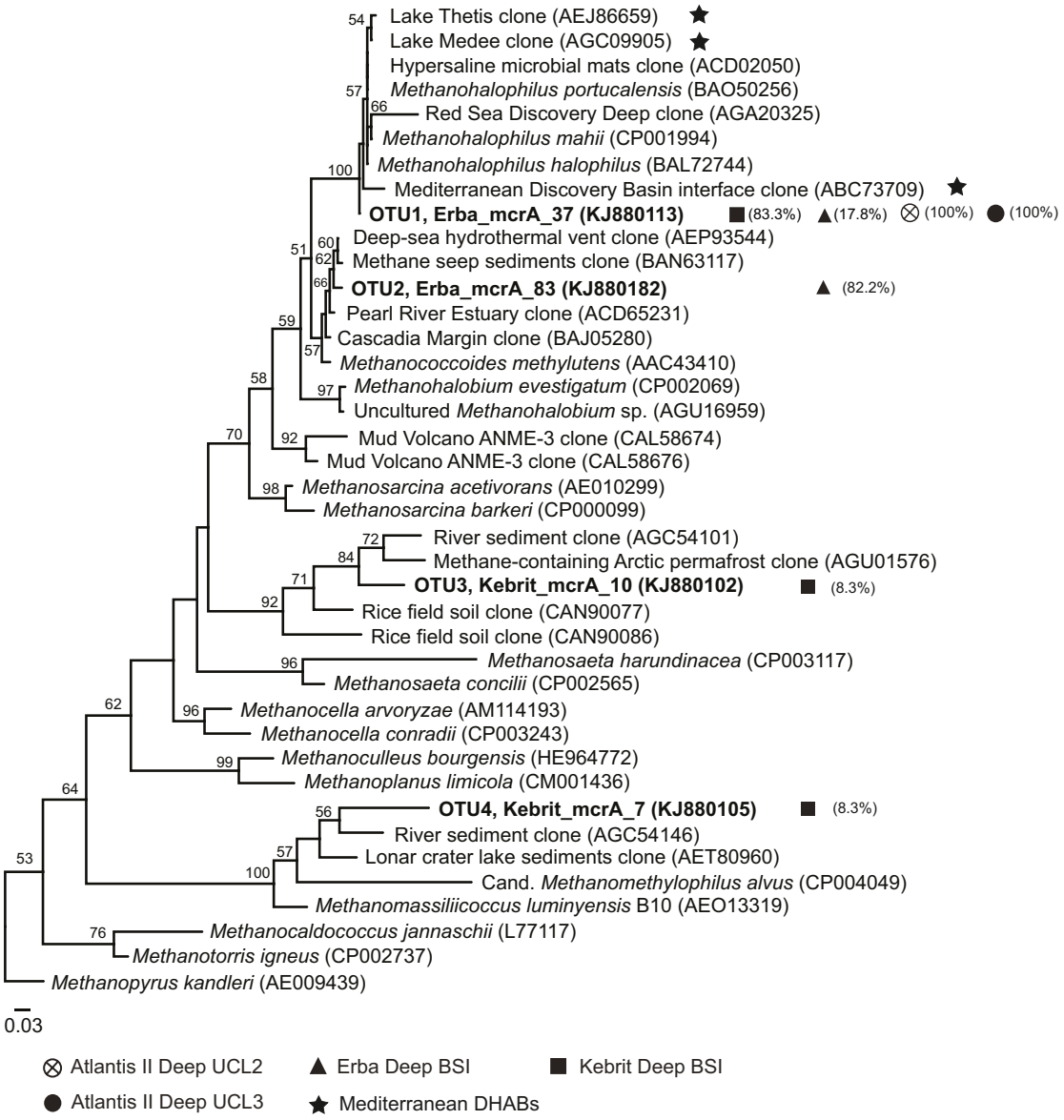


Figure 8. Phylogenetic tree of *mcrA genes* showing the relationship of representative *mcrA* clones retrieved from the deep-sea brines of the Red Sea to known methanogens and environmental sequences. Taxonomy is based on FunGene (http://fungene.cme.msu.edu). Bootstrap values are based on 1000 replicates and values above 50% are shown. Percentages in parentheses indicate the relative abundance in each sample.

Figure 8 summarizes the diversity and phylogenetic analysis based on 199 *mcrA* gene sequences retrieved from the upper convective layers of Atlantis II Deep and the BSIs of Erba and Kebrit Deep. These *mcrA gene* sequences clustered into four OTUs at a 6% amino acid sequence distance cutoff. Two

mcrA OTUs (OTU1 and OTU2, representing ~99% of the retrieved *mcrA* sequences) grouped together with cultured species of the genera *Methanohalophilus* and *Methanococcoides*, along with clones from various hypersaline environments including the deep-sea hypersaline Lake Thetis (La Cono *et al.*, 2011), Lake Medee (Yakimov *et al.*, 2013), Discovery Basin in the Mediterranean Sea (Hallsworth *et al.*, 2007), and Discovery Deep in the Red Sea (Wang *et al.*, 2013). The cultivated representatives of these genera utilize methylated compounds as methanogenic substrates and can produce methane in media with salinities of up to 4M NaCl (Boone *et al.*, 1993). This implies that the dismutation of methanol and methylamines is potentially the main methanogenic pathway in diverse hypersaline deep-sea basins. Previous studies have suggested that both methanogenesis and sulfate reduction are major energy-generating processes in the deep-sea hypersaline environment (Borin *et al.*, 2009; Daffonchio *et al.*, 2006). However, members of methanogens are in competition with sulfate-reducing bacteria for their mutual substrates (H_2 and acetate). They are also more negatively affected by increased redox potential (e.g. the increase in DO_2) and by the availability of other terminal electron acceptors (e.g., nitrate, iron, and sulfate); these conditions do exist in the brine pools. Thus, hydrogenotrophic or acetoclastic methanogens tend to be less common in hypersaline environments due to thermodynamic constraints (Scholten *et al.*, 2005). Instead, methanogens in hypersaline habitats are thought to be restricted to non-competitive substrates such as methylated amines, which occur as derivatives of compatible solutes (Ollivier *et al.*, 1994). Our results are consistent with this notion.

The remaining two rare OTUs were present only in Kebrit BSI clone library. OTU3 grouped with a previously unidentified cluster formed by clones recovered from rice field soil (Conrad *et al.*, 2008) and could not be assigned to any of the known methanogens. Interestingly, we found that OTU3 is related to *Candidatus* 'Methanoperedens nitroreducens' (Haroon *et al.*, 2013)(78% similarity) and thus could represent an anaerobic methane-oxidizing microorganism. OTU4 clustered with cultivable species of *Methanomassiliicoccales* (Borrel *et al.*, 2013), but with low amino acid sequences similarities (63%–79%). *Methanomassiliicoccales* are rarely associated with deep-sea hypersaline locations, and to our best knowledge, this is the first report of this phylotype from Red Sea brine pools, after their detection from Lake Kyros in the Mediterranean Sea (Yakimov *et al.*, 2014). Previously reported sequences associated with this order stemmed from various environments such as rumen, feces, hindguts, sludge, rice field soil, sediments, and anaerobic digesters. To date, *Methanomassiliicoccus luminyensis* is the only described species in this order (Dridi *et al.*, 2012). Based on the physiology and the genome content of this isolate, this new order utilizes an H_2-dependent methylotrophic pathway for methanogenesis (Gorlas *et al.*, 2012).

High concentrations of methane gas seem to be a common trait of most DHABs that have been studied to-date, including those studied here (Table 5). However, the link between these geochemical data and the major methane-producing taxa in such locations is still obscure. Our study showed a low diversity of methanogens in our samples and a distribution seemingly

restricted to certain layers. Several previous studies have hypothesized that members of the Candidate division MSBL1 could be the enigmatic methanogens in brine pools, given their numerical predominance among Archaea and occurrence in methane-containing layers (Yakimov *et al.*, 2013; Borin *et al.*, 2009). So far, however, there is no conclusive evidence to support this hypothesis based on the following observations. In this study, we retrieved abundant MSBL1-related 16S rRNA genes from multiple samples, but except for a single phylotype from Kebrit BSI, no novel clusters of *mcrA* genes were detected. Thus, we assume that the Candidate division MSBL1 does not possess *mcrA* genes. Still, our study does not conclusively rule out the possibility that novel *mcrA* genes have been missed due to primer bias or undersampling of the present clone libraries. Additionally, the fact that the abundance of MSBL1 16S rRNA genes seems to be correlated with locations that possess high methane concentrations might be misleading, as the methane could either be produced abiotically (Schmidt, 2003) or biotically in deeper layers of the brine pool.

1.3.5 Molecular diversity of dsrA genes

Sulfate-reducing prokaryotes are a phylogenetically diverse group of anaerobes, with the majority of them belonging to the *Deltaproteobacteria* (Muyzer and Stams, 2008). To yield energy, this group oxidizes hydrogen or small organic compounds and reduces sulfate to sulfide. The dissimilatory sulfite reductase (*dsr*), which contains two subunits (*dsrA* and *dsrB*), is the key enzyme in this process and is widely used as molecular marker to study the diversity of sulfate-reducing communities (Wagner *et al.*, 2005).

In the present study, *dsrA* clone libraries consisting of 220 sequences were constructed from the interfaces of five different brine pools in the Red Sea (Table 5). The *dsrA* gene sequences were clustered into 27 OTUs with a 6 % distance cutoff (Table S2). This is the first report on the diversity of sulfate-reducing bacteria (SRB) communities inhabiting the Red Sea brine pools. Phylogenetic analysis based on deduced amino acid sequences of *dsrA* clones revealed a high diversity of sulfate-reducing communities throughout the BSIs. Compared with the other brines in Mediterranean Sea (Yakimov *et al.*, 2014), the Red Sea brine pools seem to possess a relatively higher diversity of SRB.

The majority of *dsrA* gene sequences in Erba Deep, Kebrit Deep, and Nereus Deep were affiliated with *Desulfohalobiaceae* and the members of *Desulfobacteraceae* such as *Desulfatiglans, Desulfosalsimonas*, *Desulfobacterium,* and *Desulfobacula* (Figure 9). In contrast, in the warmer brines of Atlantis II and Discovery Deep, most of the OTUs formed distinct phylogenetic lineages clustering together with environmental sequences (Figure 9). This result indicated a different composition of sulfate-reducing communities for each geochemically distinct brine pool. In addition, the diversity of sulfate-reducing communities based on *dsrA* genes is in good agreement with the presence of sequences related to the members of *Desulfobacteraceae* and *Desulfohalobiaceae* in the 16S rRNA gene data (Figure 7).

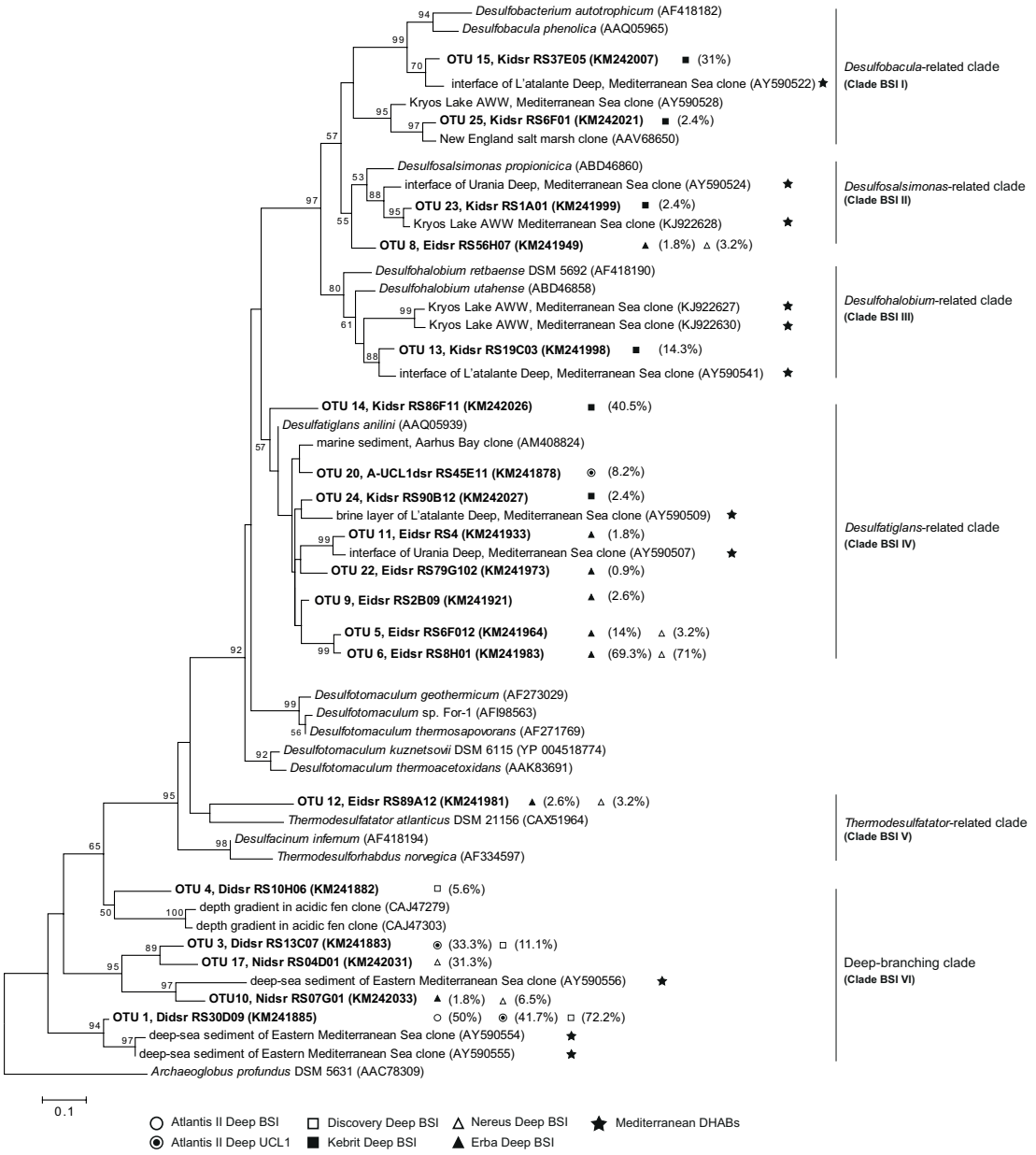


Figure 9. Phylogenetic tree based on deduced amino acid sequences of the *dsrA* clones from the brine-seawater interfaces of Red Sea brine pools, including sequences from Mediterranean DHABs. The topology of the tree is based maximum-likelihood method using 1000 bootstrap replicates. The scale bar represents 0.10 fixed mutation per nucleotide position and bootstrap values above 50% are shown. Percentages in parentheses indicate the relative abundance in each sample.

As shown in Figure 9, six different phylogenetic clades of *dsrA* gene sequences were found in the brine-seawater interfaces of the Red Sea brine pools. In Clade BSI I, two OTUs from Kebrit Deep BSI with a relative abundance of more

than 30% were affiliated with genus *Desulfobacula.* They formed clusters with the *dsrA* clones from the interface of L'Atalante Deep and Lake Kryos in the Mediterranean Sea, suggesting that the sulfate-reducing groups in these environments are specifically adapted to the conditions in the BSIs (Hallsworth *et al.*, 2007; van der Wielen and Heijs, 2007). In Clade BSI II, an OTU from Erba and one from Kebrit Deep BSI were affiliated with *Desulfosalsimonas propionicica*, a species of the *Desulfobacteraceae*. The sequences from Kebrit Deep in this clade were closely related to a clone obtained from the lower BSI of Lake Kryos (Yakimov *et al.*, 2014). The predominance of *Desulfobacteraceae* in the BSI of different brine pools might be explained by their wide range in nutritional diversity, oxygen tolerance, and metabolic plasticity.

Clade BSI III is constituted of an OTU recovered from Kebrit Deep BSI (18.2% salinity) that was distantly related to two isolated halophilic SRB species, *Desulfohalobium retbaense* (Ollivier *et al.*, 1991) and *Desulfohalobium utahense* (both growing at salinities of up to 24%) (Jakobsen *et al.*, 2006). This OTU was only present in Kebrit Deep, suggesting a preference for higher salinity environments. This finding is in a good agreement with previous analyses of *dsrAB* mRNA in the Discovery Basin of the Mediterranean Sea, which revealed that *Desulfohalobiaceae* dominated in the saltier section of the interface (from 1.60 up to 2.23 M $MgCl_2$) (Hallsworth *et al.*, 2007).

The OTUs in Clade BSI IV were widely distributed in the Red Sea brine pools (Figure 9, Table S2), especially in the ecosystems with higher sulfate concentration (Table 5). They were related to *Desulfatiglans anilini,* a sulfate-

reducing bacterium that is capable of degrading a variety of aromatic compounds including phenol (Ahn *et al.*, 2009). Almost all halophilic and halotolerant strains of SRB isolated so far are incomplete oxidizers, which oxidize organic substrates to acetate (Oren, 2011). Interestingly, sequences related to these groups are also found in Kebrit Deep (salinity 18%) and formed a cluster with a clone from the brine of L'Atalante Deep (salinity 27%), which exceeds the maximum salt limit predicted for complete oxidizers (approximately 13%) (Oren, 1999).

In Clade BSI V (Figure 9), *dsrA* gene sequences from Erba and Nereus Deep were distantly related to *Thermodesulfatator atlanticus*, a chemolithoautotrophic SRB species within the family *Thermodesulfobacteriaceae* that was isolated from a hydrothermal vent (Alain *et al.*, 2010). Some of the OTUs in the hot brines Atlantis II and Discovery Deep formed three deeply branching evolutionary lineages (Clade BSI VI) that were different from any isolated sulfate-reducing bacteria (Figure 9). As there are no cultivated representatives, the metabolism and physiology of this clade remains obscure. However, phylogenetic analysis indicates that they are related to organisms retrieved from similar environments in the Mediterranean Sea (van der Wielen and Heijs, 2007), implying that these phylotypes are specifically adapted to the DHABs. Considering the high abundance of deeply branching sequences in the Clade BSI VI (Figure 4), we assume that the interfaces of the Atlantis II and Discovery Deep harbor specific sulfate-reducing communities that are quite different to known SRB.

1.4 Conclusions

In conclusion, the bacterial communities were very diverse and, based on 16S rRNA gene copy numbers, dominated over archaea in the majority of our samples. In the multi-layered Atlantis II Deep, archaeal and bacterial communities were stratified. Marine Group I *Thaumarchaeota*, MBGE and Candidate division MSBL1 *Thermoplasmata,* halophilic methanogens, and members of class *Deltaproteobacteria* were the most common microbial groups associated with the chemoclines of the Red Sea brine pools. Methanogens were restricted to a few taxa in all studied locations, reiterating the harshness of these habitats. The sulfate-reducing communities were collectively diverse based on *dsr*A gene sequences, with the majority of the OTUs in Erba, Kebrit, and Nereus Deep being affiliated with genus *Desulfatiglans*. Additionally, the high-temperature Atlantis II and Discovery Deep harbor deeply branched lineage of *dsrA* gene sequences, which suggest that novel lineages of SRB reside in these environments. In the broader sense, these findings provide more insights on the ubiquity of methanogenic archaea and sulfate-reducers in hypersaline habitats, and should increase the impetus for future cultivation-attempts of the novel halophilic microorganisms.

1.5 Supplementary materials

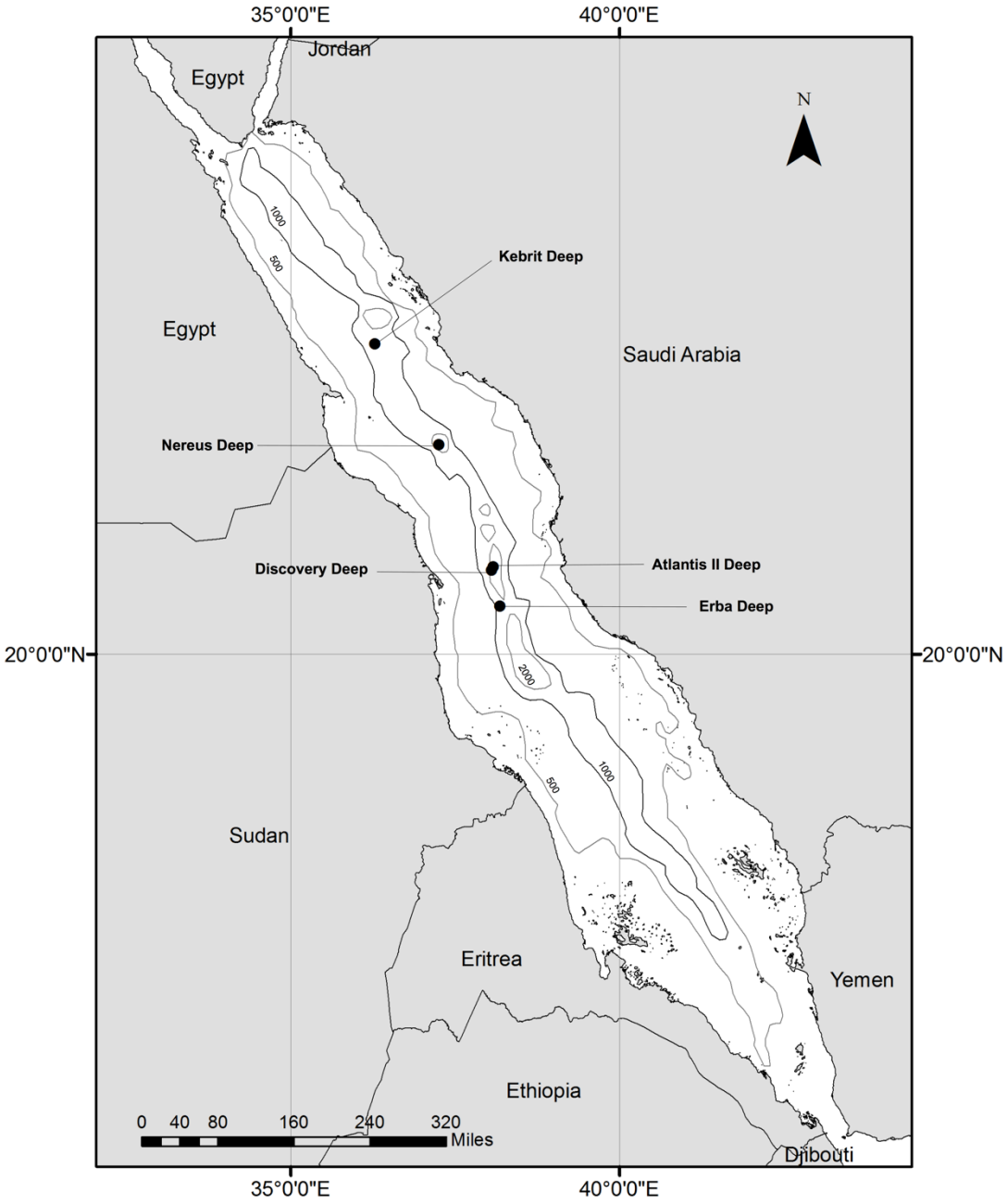


Figure S1. Sampling locations of five DHABs in the Red Sea.

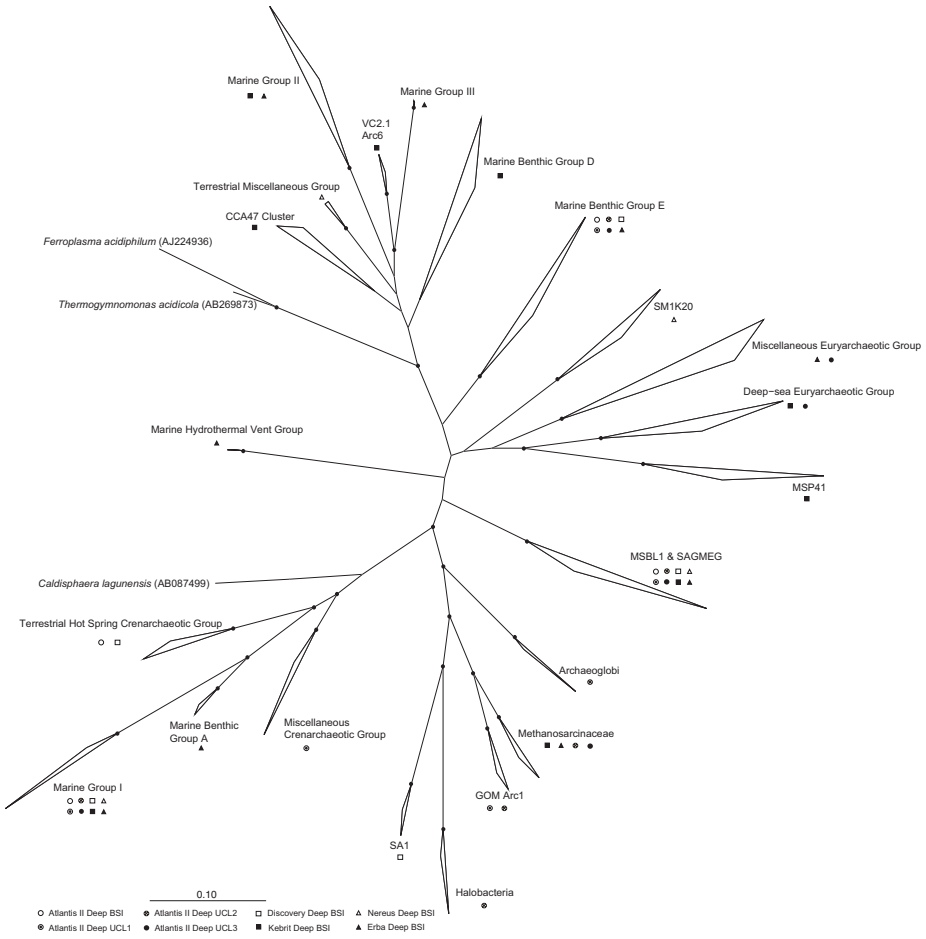


Figure S2. Phylogenetic tree showing the affiliation of archaeal lineage detected from the interfaces of the Red Sea DHABs. The tree was constructed by maximum likelihood analysis using ARB. Taxonomy is based on SSURef_115_SILVA (http://www.arb-silva.de). Dots at nodes indicate bootstrap values above 50%.

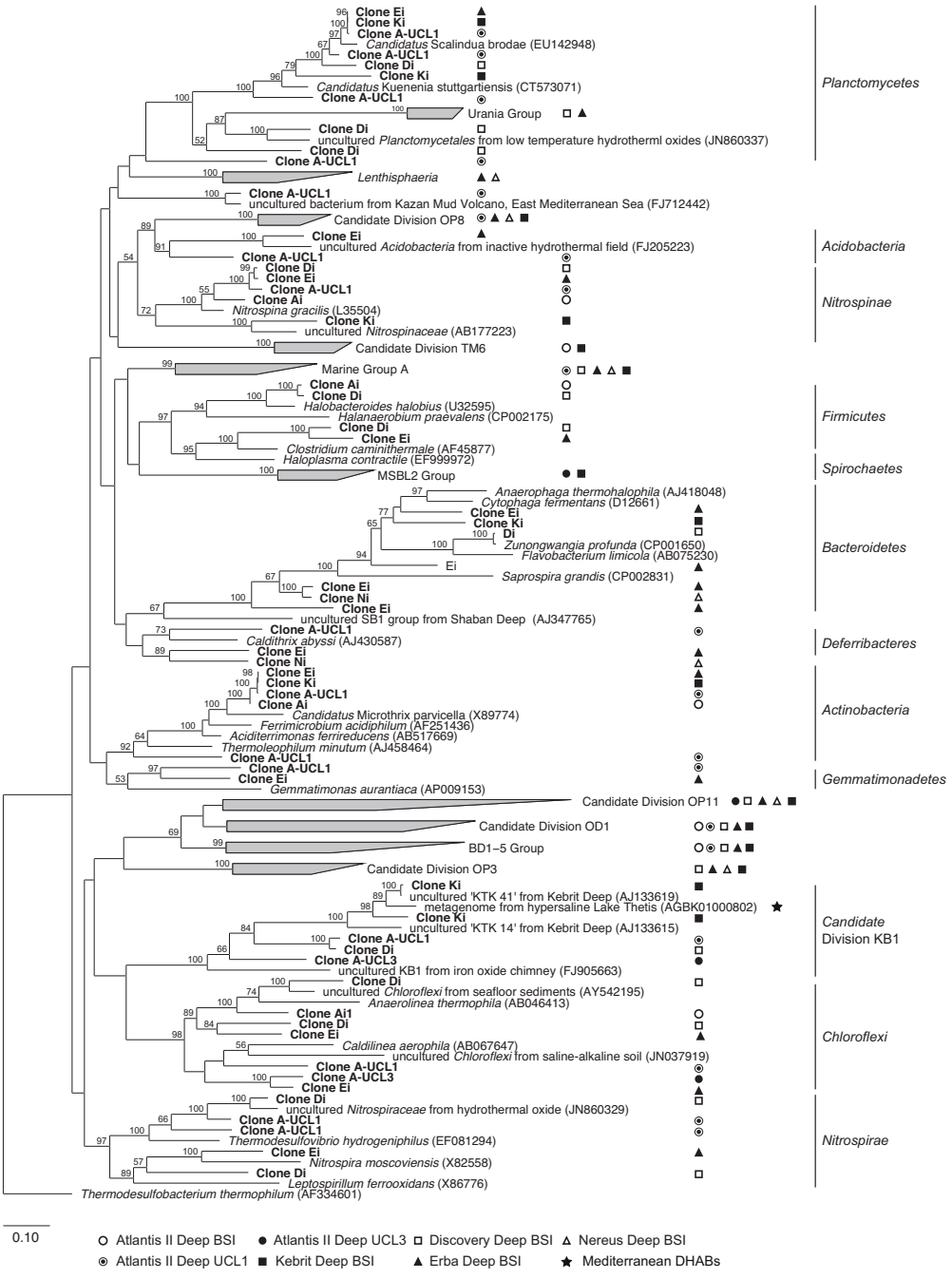


Figure S3. Major lineages of Bacteria (excluding *Proteobacteria* phylum), harboring representative sequences from the interfaces of the Red Sea DHABs. Taxonomy is based on SSURef_115_SILVA (http://www.arb-silva.de), constructed by maximum likelihood analysis using ARB.

Table S1. Shared OTUs among the sampling locations and the relative abundance of archaeal and bacterial 16S rRNA genes.

Domain	Shared OTU	Acc. no of representative sequence	Phylogenetic Affiliation	Relative abundance (%) [a]							
				Ki	Ei	Ni	Di	Ai	A-UCL1	A-UCL2	A-UCL3
Archaea	OTU1	KJ881931	Marine Group I	68.8	67.7	70.2	76.3	85.3	9.4	13.0	17.1
	OTU2	KJ881537	Marine Benthic Group E	-	1.3	-	-	1.1	77.1	48.0	8.6
	OTU3	KJ882098	MSBL1	9.4	-	14.3	-	-	-	-	-
	OTU4	KJ881658	Marine Benthic Group E	-	-	-	1.7	4.2	13.5	-	-
	OTU5	KJ881514	Marine Group I	-	3.9	4.8	-	2.1	-	-	11.4
	OTU7	KJ881716	MSBL1	1.3	-	-	-	-	-	8.0	2.9
	OTU8	KJ881678	*Methanosarcinaceae*	2.5	-	-	-	-	-	3.0	11.4
	OTU9	KJ881876	MSBL1	-	-	-	7.6	2.1	-	-	-
	OTU10	KJ881874	Terrestrial Hot Spring Group	-	-	-	3.4	6.3	-	-	-
	OTU11	KJ881571	GOM Arc1	-	-	-	-	-	5.2	2.0	-
	OTU12	KJ881653	South African Goldmine Euryarchaeotal Group	-	-	-	-	-	-	3.0	5.7
	OTU14	KJ881974	Marine Group I	1.3	1.9	-	-	-	-	-	-
	OTU40	KJ881892	Marine Group I	-	0.6	1.2	-	-	-	-	-
	OTU46	KJ882127	MSBL1	0.6	-	1.2	-	-	-	-	-
Bacteria	OTU3	KM018757	*Desulfobacteraceae*	-	0.2	1	2.6	-	-	-	-
	OTU5	KM018689	*Halomonadaceae*	-	-	-	24.3	1.1	-	-	-
	OTU7	KM019134	SAR406	-	0.1	1	0.9	-	0.6	-	-
	OTU12	KM018842	*Desulfobacteraceae*	-	0.1	1	1.7	-	-	-	-
	OTU13	KM018808	*Candidatus Scalindua*	0.7	0.1	-	-	-	3.4	-	-
	OTU14	KM018800	*Desulfobacteraceae*	3.9	0.1	-	-	-	1.1	-	-
	OTU17	KM018890	*Gammaproteobacteria* E01-9C-26 group	2	0.1	-	-	-	1.1	-	-
	OTU19	KM018941	*Acidimicrobineae*	2	0.1	-	-	2.2	0.6	-	-
	OTU21	KM018785	Candidate division OP3	-	0.1	1.9	-	-	-	-	-
	OTU23	KM018927	Candidate division OP8	-	0.1	1	-	-	0.6	-	-
	OTU24	KM018592	*Nitrospinaceae*	-	-	-	-	1.1	3.4	-	-
	OTU28	KM018812	Candidate division OP3	-	0.1	-	0.9	-	-	-	-
	OTU29	KM018950	SAR324	-	0.1	1	0.9	-	-	-	-
	OTU32	KM018372	*Mariprofundaceae*	-	-	-	-	5.4	0.6	-	-
	OTU33	KM018974	SAR11	0.7	0.1	-	2.6	-	-	-	-
	OTU34	KM018857	SAR406	-	0.1	-	0.9	-	-	-	-
	OTU35	KM018734	*Alteromonadaceae*	-	-	-	3.5	-	0.6	-	-
	OTU36	KM018506	*Nitrospinaceae*	-	-	-	0.9	-	2.3	-	-
	OTU54	KM018721	*Nitrospinaceae*	0.7	0.1	-	0.9	-	-	-	-
	OTU56	KM018968	*Phycisphaeraceae*	-	0.1	-	0.9	-	-	-	-
	OTU57	KM018934	*Sphingobacteriales*	-	0.1	1	-	-	-	-	-
	OTU77	KM018692	*Halomonadaceae*	-	-	-	0.9	-	1.1	-	-
	OTU87	KM018815	SAR406	-	0.1	-	0.9	-	-	-	-
	OTU97	KM019101	SAR324	0.7	0.1	-	-	-	-	-	-

[a]The calculation of relative abundance is based on the number of clones in each OTU divided by total clones per location.

Table S2. OTU classification and the relative abundance of *dsrA* genes in each sampling locations.

OTUs	Relative abundance (%)[a]					
	Ki	**Ei**	**Ni**	**Di**	**Ai**	**A-UCL1**
OTU1				**72.22**	**50**	**41.67**
OTU2				11.11	25	8.33
OTU3				11.11		33.33
OTU4				5.55		
OTU5		14.0	3.2			
OTU6		**69.3**	71.0			
OTU7		2.6	3.2			
OTU8		1.8	3.2			
OTU9		2.6				
OTU10		1.8	6.5			
OTU11		1.8				
OTU12		2.6	3.2			
OTU13	14.3					
OTU14	**40.5**					
OTU15	31.0					
OTU16	4.8					
OTU17			6.5			
OTU18					25	
OTU19						8.33
OTU20						8.33
OTU21		0.9				
OTU22		0.9				
OTU23	2.4					
OTU24	2.4					
OTU25	2.4					
OTU26			3.2			
OTU27	2.4					

[a]The calculation of relative abundance is based on the number of clones in each OTU divided by total clones per location.

1.6 References

Abdallah RZ, Adel M, Ouf A, Sayed A, Ghazy MA, Alam I, *et al.* (2014). Aerobic methanotrophic communities at the Red Sea brine-seawater interface. *Front Microbiol* **5**. doi:10.3389/fmicb.2014.00487.

Ahn Y-B, Chae J-C, Zylstra GJ, Haeggblom MM. (2009). Degradation of phenol via phenylphosphate and carboxylation to 4-hydroxybenzoate by a newly isolated strain of the sulfate-reducing bacterium *Desulfobacterium anilini*. *Appl Environ Microbiol* **75**: 4248–4253.

Alain K, Postec A, Grinsard E, Lesongeur F, Prieur D, Godfroy A. (2010). *Thermodesulfatator atlanticus* sp. nov., a thermophilic, chemolithoautotrophic, sulfate-reducing bacterium isolated from a Mid-Atlantic Ridge hydrothermal vent. *Int J Syst Evol Microbiol* **60**: 33–38.

Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, *et al.* (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* **25**: 3389–3402.

Angel R, Matthies D, Conrad R. (2011). Activation of methanogenesis in arid biological soil crusts despite the presence of oxygen. Gilbert, JA (ed). *PLoS ONE* **6**: e20453.

Antunes A, Eder W, Fareleira P, Santos H, Huber R. (2003). *Salinisphaera shabanensis* gen. nov., sp. nov., a novel, moderately halophilic bacterium from the brine-seawater interface of the Shaban Deep, Red Sea. *Extremophiles* **7**: 29–34.

Antunes A, França L, Rainey FA, Huber R, Nobre MF, Edwards KJ, *et al.* (2007). *Marinobacter salsuginis* sp. nov., isolated from the brine-seawater interface of the Shaban Deep, Red Sea. *Int J Syst Evol Microbiol* **57**: 1035–1040.

Antunes A, Ngugi DK, Stingl U. (2011). Microbiology of the Red Sea (and other) deep-sea anoxic brine lakes. *Environmental Microbiology Reports* **3**: 416–433.

Backer H, Schoell M. (1972). New deeps with brines and metalliferous sediments in the Red Sea. *Nat Phys Sci* **240**: 153–158.

Boone DR, Mathrani IM, Liu Y, Menaia JAGF, Mah RA, Boone JE. (1993). Isolation and characterization of *Methanohalophilus portucalensis* sp. nov. and DNA reassociation study of the genus *Methanohalophilus*. *Int J Syst Bacteriol* **43**: 430–437.

Borin S, Brusetti L, Mapelli F, D'Auria G, Brusa T, Marzorati M, *et al.* (2009). Sulfur cycling and methanogenesis primarily drive microbial colonization of the highly sulfidic Urania deep hypersaline basin. *Proc Natl Acad Sci USA* **106**: 9151–9156.

Borrel G, O'Toole PW, Harris HMB, Peyret P, Brugere JF, Gribaldo S. (2013). Phylogenomic data support a seventh order of methylotrophic methanogens and provide insights into the evolution of methanogenesis. *Genome Biol Evol* **5**: 1769–1780.

Bougouffa S, Yang JK, Lee OO, Wang Y, Batang Z, Al-Suwailem A, *et al.* (2013). Distinctive microbial community structure in highly stratified deep-sea brine water columns. *Appl Environ Microbiol* **79**: 3425–3437.

Brune A, Frenzel P, Cypionka H. (2000). Life at the oxic-anoxic interface: microbial activities and adaptations. *FEMS Microbiology Reviews* **24**: 691–710.

Conrad R, Klose M, Noll M, Kemnitz D, Bodelier PLE. (2008). Soil type links microbial colonization of rice roots to methane emission. *Global Change Biology* **14**: 657–669.

Daffonchio D, Borin S, Brusa T, Brusetti L, van der Wielen PWJJ, Bolhuis H, *et al.* (2006). Stratified prokaryote network in the oxic-anoxic transition of a deep-sea halocline. *Nature* **440**: 203–207.

Dridi B, Fardeau M-L, Ollivier B, Raoult D, Drancourt M. (2012). *Methanomassiliicoccus luminyensis* gen. nov., sp nov., a methanogenic archaeon isolated from human faeces. *Int J Syst Evol Microbiol* **62**: 1902–1907.

Eder W, Jahnke LL, Schmidt M, Huber R. (2001). Microbial diversity of the brine-seawater interface of the Kebrit Deep, Red Sea, studied via 16S rRNA gene sequences and cultivation methods. *Appl Environ Microbiol* **67**: 3077–3085.

Eder W, Ludwig W, Huber R. (1999). Novel 16S rRNA gene sequences retrieved from highly saline brine sediments of kebrit deep, red Sea. *Arch Microbiol* **172**: 213–218.

Eder W, Schmidt M, Koch M, Garbe-Schönberg D, Huber R. (2002). Prokaryotic phylogenetic diversity and corresponding geochemical data of the brine-seawater interface of the Shaban Deep, Red Sea. *Environ Microbiol* **4**: 758–763.

Edgar RC, Haas BJ, Clemente JC, Quince C, Knight R. (2011). UCHIME improves sensitivity and speed of chimera detection. *Bioinformatics* **27**: 2194–2200.

Ferry JG. (2001). Methanogenesis Biochemistry. John Wiley & Sons, Ltd: Chichester, UK doi:10.1038/npg.els.0000573.

Fish JA, Chai B, Wang Q, Sun Y, Brown CT, Tiedje JM, *et al.* (2013). FunGene: the functional gene pipeline and repository. *Front Microbiol* **4**. doi:10.3389/fmicb.2013.00291.

Friedrich MW. (2005). Methyl-coenzyme M reductase genes: unique functional markers for methanogenic and anaerobic methane-oxidizing archaea. *Meth Enzymol* **397**: 428–442.

Gorlas A, Robert C, Gimenez G, Drancourt M, Raoult D. (2012). Complete genome sequence of *Methanomassiliicoccus luminyensis*, the largest genome of a human-associated *Archaea* species. **194**: 4745–4745.

Hallsworth JE, Yakimov MM, Golyshin PN, Gillion JLM, D'Auria G, de Lima Alves F, *et al.* (2007). Limits of life in $MgCl_2$-containing environments: chaotropicity defines the window. *Environ Microbiol* **9**: 801–813.

Haroon MF, Hu S, Shi Y, Imelfort M, Keller J, Hugenholtz P, *et al.* (2013). Anaerobic oxidation of methane coupled to nitrate reduction in a novel archaeal lineage. *Nature* **500**: 567–570.

Hartmann M, Scholten JC, Stoffers P, Wehner F. (1998). Hydrographic structure of brine-filled deeps in the Red Sea - new results from the Shaban, Kebrit, Atlantis II, and Discovery Deep. *Marine Geology* **144**: 311–330.

Jakobsen TF, Kjeldsen KU, Ingvorsen K. (2006). Desulfohalobium utahense sp. nov., a moderately halophilic, sulfate-reducing bacterium isolated from Great Salt Lake. *Int J Syst Evol Microbiol* **56**: 2063–2069.

Kato S, Kobayashi C, Kakegawa T, Yamagishi A. (2009). Microbial communities in iron-silica-rich microbial mats at deep-sea hydrothermal fields of the Southern Mariana Trough. *Environ Microbiol* **11**: 2094–2111.

Kato S, Takano Y, Kakegawa T, Oba H, Inoue K, Kobayashi C, *et al.* (2010). Biogeography and biodiversity in sulfide structures of active and inactive vents at eeep-sea hydrothermal fields of the southern Mariana Trough. *Appl Environ Microbiol* **76**: 2968–2979.

Kendall MM, Wardlaw GD, Tang CF, Bonin AS, Liu Y, Valentine DL. (2007). Diversity of *Archaea* in marine sediments from Skan Bay, Alaska, including cultivated methanogens, and description of *Methanogenium boonei* sp. nov. *Appl Environ Microbiol* **73**: 407–414.

Klindworth A, Pruesse E, Schweer T, Peplies J, Quast C, Horn M, *et al.* (2013). Evaluation of general 16S ribosomal RNA gene PCR primers for classical and next-generation sequencing-based diversity studies. *Nucleic Acids Res* **41**: e1–e1.

Konneke M, Bernhard AE, la Torre de JR, Walker CB, Waterbury JB, Stahl DA. (2005). Isolation of an autotrophic ammonia-oxidizing marine archaeon. *Nature* **437**: 543–546.

La Cono V, Smedile F, Bortoluzzi G, Arcadi E, Maimone G, Messina E, *et al.* (2011). Unveiling microbial life in new deep-sea hypersaline Lake Thetis. Part I: Prokaryotes and environmental settings. *Environ Microbiol* **13**: 2250–2268.

Lane DJ. (1991). 16S/23S rRNA sequencing. *Nucleic Acid Techniques in Bacterial Systematics*

125–175.

LaRock PA, Lauer RD, Schwarz JR, Watanabe KK, Wiesenburg DA. (1979). Microbial biomass and activity distribution in an anoxic, hypersaline basin. *Appl Environ Microbiol* **37**: 466–470.

Ludwig W, Strunk O, Westram R, Richter L, Meier H, Yadhukumar, *et al.* (2004). ARB: a software environment for sequence data. *Nucleic Acids Res* **32**: 1363–1371.

Marschall C, Frenzel P, Cypionka H. (1993). Influence of oxygen on sulfate reduction and growth of sulfate-reducing bacteria. *Arch Microbiol* **159**: 168–173.

Muyzer G, Stams AJM. (2008). The ecology and biotechnology of sulphate-reducing bacteria. *Nature Reviews Microbiology* **6**: 441–454.

Ngugi DK, Blom J, Alam I, Rashid M, Ba-Alawi W, Zhang G, *et al.* (2014). Comparative genomics reveals adaptations of a halotolerant thaumarchaeon in the interfaces of brine pools in the Red Sea. *ISME J*. doi:10.1038/ismej.2014.137.

Ngugi DK, Stingl U. (2012). Combined analyses of the ITS loci and the corresponding 16S rRNA genes reveal high micro- and macrodiversity of SAR11 populations in the Red Sea. Rodriguez-Valera, F (ed). *PLoS ONE* **7**: e50274.

Ollivier B, Caumette P, Garcia JL, Mah RA. (1994). Anaerobic bacteria from hypersaline environments. *Microbiol Rev* **58**: 27–38.

Ollivier B, Hatchikian CE, Prensier G, GUEZENNEC J, Garcia JL. (1991). *Desulfohalobium retbaense* gen. nov., sp. nov., a Halophilic Sulfate-Reducing Bacterium from Sediments of a Hypersaline Lake in Senegal. *Int J Syst Bacteriol* **41**: 74–81.

Oren A. (1999). Bioenergetic aspects of halophilism. *Microbiol Mol Biol Rev* **63**: 334–348.

Oren A. (2011). Diversity of Halophiles. In: *Extremophiles Handbook*, Springer Japan: Tokyo, pp 309–325.

Park B-J, Park S-J, Yoon D-N, Schouten S, Sinninghe Damsté JS, Rhee S-K. (2010). Cultivation of autotrophic ammonia-oxidizing archaea from marine sediments in coculture with sulfur-oxidizing bacteria. *Appl Environ Microbiol* **76**: 7575–7587.

Pester M, Bittner N, Deevong P, Wagner M, Loy A. (2010). A 'rare biosphere' microorganism contributes to sulfate reduction in a peatland. *ISME J* **4**: 1591–1602.

Pruesse E, Quast C, Knittel K, Fuchs BM, Ludwig W, Peplies J, *et al.* (2007). SILVA: a comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB. *Nucleic Acids Res* **35**: 7188–7196.

Sass AM, Sass H, Coolen MJL, Cypionka H, Overmann J. (2001). Microbial communities in the chemocline of a hypersaline deep-sea basin (Urania basin, Mediterranean Sea).

Schloss PD, Westcott SL, Ryabin T, Hall JR, Hartmann M, Hollister EB, *et al.* (2009). Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Appl Environ Microbiol* **75**: 7537–7541.

Schmidt M. (2003). High-resolution methane profiles across anoxic brine–seawater boundaries in the Atlantis-II, Discovery, and Kebrit Deeps (Red Sea). *Chemical Geology* **200**: 359–375.

Scholten J, Joye SB, Hollibaugh JT, Murrell JC. (2005). Molecular analysis of the sulfate reducing and archaeal community in a meromictic soda lake (Mono Lake, California) by targeting 16S rRNA, mcrA, apsA, and dsrAB genes. *Microb Ecol* **50**: 29–39.

Siam R, Mustafa GA, Sharaf H, Moustafa A, Ramadan AR, Antunes A, *et al.* (2012). Unique prokaryotic consortia in geochemically distinct sediments from Red Sea Atlantis II and Discovery Deep brine pools Mormile, MR (ed). *PLoS ONE* **7**. doi:10.1371/journal.pone.0042872.

Springer E, Sachs MS, Woese CR, Boone DR. (1995). Partial gene sequences for the A subunit of methyl-coenzyme M reductase (mcrI) as a phylogenetic tool for the family

Methanosarcinaceae. *Int J Syst Bacteriol* **45**: 554–559.

Swan BK, Chaffin MD, Martinez-Garcia M, Morrison HG, Field EK, Poulton NJ, *et al.* (2014). Genomic and Metabolic Diversity of Marine Group I Thaumarchaeota in the Mesopelagic of Two Subtropical Gyres Randau, L (ed). *PLoS ONE* **9**. doi:10.1371/journal.pone.0095380.

Takai K, Nunoura T, Ishibashi J-I, Lupton J, Suzuki R, Hamasaki H, *et al.* (2008). Variability in the microbial communities and hydrothermal fluid chemistry at the newly discovered Mariner hydrothermal field, southern Lau Basin. *Journal of Geophysical Research-Biogeosciences* **113**. doi:10.1029/2007JG000636.

Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S. (2011). MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol Biol Evol* **28**: 2731–2739.

Thompson JD, Higgins DG, Gibson TJ. (1994). Clustal W: Improving the Sensitivity of Progressive Multiple Sequence Alignment Through Sequence Weighting, Position-Specific Gap Penalties and Weight Matrix Choice. *Nucleic Acids Res* **22**: 4673–4680.

van der Wielen PWJJ. (2005). The enigma of prokaryotic life in deep hypersaline anoxic basins. *Science* **307**: 121–123.

van der Wielen PWJJ, Heijs SK. (2007). Sulfate-reducing prokaryotic communities in two deep hypersaline anoxic basins in the Eastern Mediterranean deep sea. *Environ Microbiol* **9**: 1335–1340.

Wagner M, Loy A, Klein M, Lee N, Ramsing NB, Stahl DA, *et al.* (2005). Functional marker genes for identification of sulfate-reducing prokaryotes. *Meth Enzymol* **397**: 469–489.

Wang Y, Cao H, Zhang G, Bougouffa S, Lee OO, Al-Suwailem A, *et al.* (2013). Autotrophic microbe metagenomes and metabolic pathways differentiate adjacent red sea brine pools. *Sci Rep* **3**: 1748.

Wang Y, Yang J, Lee OO, Dash S, Lau SCK, Al-Suwailem A, *et al.* (2011). Hydrothermally generated aromatic compounds are consumed by bacteria colonizing in Atlantis II Deep of the Red Sea. *ISME J* **5**: 1652–1659.

Yakimov MM, Giuliano L, Cappello S, Denaro R, Golyshin PN. (2007). Microbial community of a hydrothermal mud vent underneath the deep-sea anoxic brine lake Urania (eastern Mediterranean). *Orig Life Evol Biosph* **37**: 177–188.

Yakimov MM, La Cono V, Denaro R, D'Auria G, Decembrini F, Timmis KN, *et al.* (2007). Primary producing prokaryotic communities of brine, interface and seawater above the halocline of deep anoxic lake L'Atalante, Eastern Mediterranean Sea. *ISME J* **1**: 743–755.

Yakimov MM, La Cono V, La Spada G, Bortoluzzi G, Messina E, Smedile F, *et al.* (2014). Microbial community of the deep-sea brine Lake Kryos seawater-brine interface is active below the chaotropicity limit of life as revealed by recovery of mRNA. *Environ Microbiol* **17**: n/a–n/a.

Yakimov MM, La Cono V, Slepak VZ, La Spada G, Arcadi E, Messina E, *et al.* (2013). Microbial life in the Lake Medee, the largest deep-sea salt-saturated formation. **3**. doi:10.1038/srep03554.

2 Chapter II GENOMIC AND TRANSCRIPTOMIC INSIGHTS INTO THE GENUS METHANOHALOBIUM WITH A NEW METHANOGEN FROM THE ERBA DEEP, RED SEA

Yue Guan[a], David K. Ngugi[a], Manikandan Vinu[a], Craig Michell[a], Jochen Blom[b], Sylvain Guillot[a], James G. Ferry[c], Ulrich Stingl[a]

[a]Red Sea Research Center, King Abdullah University of Science and Technology, Thuwal, Saudi Arabia

[b]Justus-Liebig-Universität Giessen, Bioinformatik und Systembiologie, Giessen, Germany

[c]Department of Biochemistry and Molecular Biology, Pennsylvania State University, University Park, Pennsylvania, USA

Author contributions:

Yue Guan isolated and characterized *Methanohalobium* strain ERBA, analyzed the data, and wrote the manuscript. Yue Guan, David K. Ngugi, Manikandan Vinu performed the bioinformatic analyses. Craig Michell and Yue Guan performed RNA-seq. Sylvain Guillot assisted with the characterization. Jochen Blom developed and maintained bioinformatic platforms used in this work. Ulrich Stingl and James G. Ferry provided directions and lab support for the study.

Abstract

Members of methanogenic archaea can grow under a wide range of salinities. The ones that can tolerate extremely high salt concentrations are rare. Here, we introduce the second *Methanohalobium* species, designated as strain ERBA isolated from Erba Deep of the Red Sea. This is the first report of a new methanogenic isolate from the deep-sea brines of the Red Sea. This additional species has made a comparative genomic study of the *Methanohalobium* genus possible. By comparing these two closely related halophilic methanogen species, we reveal intensive genomic capacities for energy and carbohydrate metabolisms, and diverse strategies for static and dynamic osmoadaptation. Having unique genes for energy and central metabolisms, membrane transport and cell membrane signatures, owning the coding capabilities to resist toxic compounds and foreign DNA element infections might have given the deep-sea brine isolate strain ERBA advantages over other methanogens to succeed in its extreme habitat. In addition, we used RNA-seq methodologies to determine differentially expressed genes of strain ERBA under low, moderate, and elevated salt concentrations. Our data revealed that this deep-sea halophilic methanogen could utilize diverse metabolisms and pathways to combat unfavorable salt conditions. Among other factors, the upregulation of genes correlated with inorganic ion transport and osmoadaptation, stringent control of amino acid and protein synthesis and the ability to regulate energy metabolism and electron transfer might contribute to its adaptation to less desirable salt concentrations.

2.1 Introduction

Methane is a green house gas and can be utilized as an alternative energy source. Methanogens predominantly contribute to biological generation of methane (Ferry, 1994). Methanogenesis plays an important role in the global carbon cycle (reviewed in (Thauer *et al.*, 2008)). Halophilic methane-producing archaea are limited primarily to hypersaline and oxygen-starved habitats (McGenity, 2010). The genus *Methanohalobium* is the only extreme halophilic clade of methane-producing archaea with so far only one cultivated and described species, *Methanohalobium evestigatum* (Zhilina and Zavarzin, 1987). Additional isolates of methanogens within this clade and additional genomic sequences will not only enable comparative analyses of extreme halophilic methanogens but also provide insights into genomic features and adaptations of halophilic methanogens.

Deep-sea brines of the Red Sea are among the most challenging, remote and hypersaline polyextreme habitats on the planet Earth. The Red Sea harbors a remarkable number (approx. 25) of these deep-sea brines along its rift valley region, a divergent plate boundary where the Asian and African plates drift apart (Antunes *et al.*, 2011). They provide an outstanding opportunity to understand how life could adapt to deep and hypersaline ecosystems. In our recent study on the methanogenic and sulfate reducing microbial communities in five of the Red Sea deep-sea brines (also described in Chapter I (Guan *et al.*, 2015)), we have identified a 16S rRNA sequence that is closely related to cultivated *Methanohalobium evestigatum* (98% identity). As subsequently described in this chapter, we successfully isolated a new *Methanohalobium*

strain from one of the deep-sea brine-pools, Erba Deep, in the Red Sea and obtained its draft genome sequence. We furthermore explore the genome dataset of both strains and describe the conserved and variable genomic traits of this clade to highlight their diverse and unique strategies for energy conservation and salt adaption.

Salinity is one of the most important physiochemical constraints limiting methanogenesis (Hoehler *et al.*, 2010). However, since no transcriptomics study has ever been done on halophilic methane producing archaea before, little is known about the global transcriptomic characteristics, in particular those underlying the responsive mechanisms of halophilic methanogens to environmental variables. To examine the effect of salinity, especially at high salt concentrations, to gene expressions of *Methanohalobium*, we performed a comparative transcriptomic investigation on cells of the strain ERBA growing across three different salinities to examine their adaptation to salt at the RNA level. The goal of this chapter is to provide a strong genomic and transcriptomic base to further our understanding and knowledge on how methanogens cope with salt and polyextreme environments.

2.2 Methods

2.2.1 Sampling

In November 2011 during the R/V Aegaeo cruise to the Red Sea, the brine-seawater interface of the Erba Deep (20° 44' 4" N, 38° 10' 50" E) was sampled by a Rosette Sampler at a depth of 2381 meters below the sea surface as previously described (Guan *et al.*, 2015; Ngugi *et al.*, 2014).

2.2.2 Isolation and characterization of methanogens

Using the water sample collected from Erba Deep as an inoculum, anaerobic enrichment of methane-producing cultures followed by subsequent direct plating with agar overlays under strictly anaerobic conditions was conducted to obtain the pure culture of the strain ERBA. Sterile media preparation and formula was modified from the High Salt medium described by Sowers and Ferry et al. on isolation of marine methanogens *Methanococcoides methylutens* and *Methanosarcina acetivorans* (Sowers and Ferry, 1983; Sowers *et al.*, 1984) by increasing the salinity gradient to target halophilic methanogens. 50mM trimethylamine (TMA), dimethylamine (DMA) and monomethylamine (MMA), 150 mM methanol, 90mM acetate, 50mM formate and H_2-CO_2 (1:4 vol/vol) were supplemented separately to test for potential substrate utilization. The effects of different environmental parameters (temperature, salinity and pH) were tested. Cell growth was monitored by measuring optical density at OD_{600} under a SpectraMax® Multi-Mode Microplate Reader (Molecular Devices). Methane production was measured by repeated removal from the sample of 0.25 ml of the gas phase with a Pressure-lok® precision analytical syringe (Sigma). The samples were subsequently analyzed by an Agilent 6890 Gas Chromatograph (Agilent).

2.2.3 Microscopy

For visualization using cryo-scanning electron microscopy(cryo-SEM), cells grown to exponential phase were fixed in 2.5% glutaraldehyde. Micrographs were obtained at KAUST's Advanced Nanofabrication and Image Core Laboratory by using a Quorum PP2000T cryo-transfer system (Qurorum

Technologies, Newhaven, UK) attached to an FEI Nova Nano630 SEM with a field emission electron source and through-the-lens electron detectors.

2.2.4 Whole-genome sequencing, assembly, annotation, and comparative genomics

Genomic DNA of strain ERBA was extracted using a DNeasy Blood & Tissue Kit (Qiagen). The genome of *Methanohalobium* strain ERBA was sequenced at the Bioscience Core Laboratory of King Abdullah University of Science and Technology (KAUST) using the Illumina MiSeq technology. SPAdes Genome Assembler was used for the *de novo* draft genome assembly and error correction (Bankevich *et al.*, 2012). Genes were identified and annotated by both the INDIGO pipeline (Alam *et al.*, 2013) and Integrated Microbial Genomes (IMG), and followed by a round of manual curation. The completeness of the genome was determined via AmphoraNet based on the presence or absence of 104 single-copy genes in Archaea (Kerepesi *et al.*, 2014). The chromosomal origin (oriC) regions of each genome were predicted using Ori-Finder2 (http://tubic.tju.edu.cn/Ori-Finder2/). Clusters of Orthologous Groups (COG) classification and Pfam classification of predicted proteins was done by WebMGA (http://weizhong-lab.ucsd.edu/metagenomic-analysis/, last accessed Oct 2015) (Sitao Wu *et al.*, 2011) at an E value of 10^{-5} for the genomes. CRISPR loci were detected using CRISPR Finder (Grissa *et al.*, 2007) and compared using CRISPRcompar (Grissa *et al.*, 2008) and CRISPRmap v2.1-2014 (Lange *et al.*, 2013). The latter contains 4719 consensus repeats covering 24 families and 18 structural motifs. Prophage screening was done using PHAST (Zhou *et al.*, 2011).

To evaluate the inter-species relationship, average nucleotide identity (ANI) based on the BLAST algorithm, average amino acid identity (AAI) and conservation of gene order (synteny) were analyzed by JSpecies (Richter and Rosselló-Mora, 2009) and a synteny script described previously (Yelton *et al.*, 2011). To classify the core genome and variable genes of *Methanohalobium* species, the *Methanohalobium evestigatum* chromosome was re-annotated using the INDIGO pipeline (Alam *et al.*, 2013) harmonized with the genome annotation of strain ERBA. Orthologues were identified by a BLAST Score Ratio (BSR) approach with a threshold value of 0.4 that represents *c.* 30% of amino acid identity over *c.* 30% peptide length (Rasko *et al.*, 2005) on the EDGAR platform (Blom *et al.*, 2009). To identify low-synteny loci and genome rearrangement in both genomes, Progressive MAUVE was used to perform whole-genome sequence (Darling *et al.*, 2009).

2.2.5 Phylogenetic analysis

Genome sequences were downloaded from the NCBI FTP site (Available online: ftp://ftp.ncbi.nlm.nih.gov/genomes/Bacteria/). A total of 77 AMPHORA2-selected marker genes (Martin Wu and Scott, 2012) were identified to be present in all seven-order methanogenic archaea as well as a few selected *Thermoplasmata* genomes selected for the phylogenetic analysis. The phylogeny of these methane-producing archaea was inferred by a maximum likelihood approach with Le Gascuel's substitution model under 100 bootstrap replicates implemented in RAxML (Stamatakis, 2006).

2.2.6 RNA extraction and RNA-seq

An overall schematic diagram of methods used in the RNA-seq part of this chapter is presented in Figure S1 in the supplementary materials.

For comparative transcriptomics analysis, cells of strain ERBA were grown in a TMA based medium under 3 different salinity (NaCl) conditions: 5% (low salt), 10% (optimal growth salinity) and 15% (high salt) with three independent replicates under each growth condition. Cells were harvested at early log-phase (OD_{600} = 0.2) through centrifugation (10,000x g, 5 minutes). The cell pellets were lysed by TRIzol® Reagent (Invitrogen, USA) followed by purification procedures. The total RNA quality and quantity were measured with BioAnalyzer 2100 Total RNA Nano Kit (Agilent Technologies Inc., USA). All extracted total RNA samples have Ribosomal Integrity Numbers (RIN) greater than 9. rRNAs were subsequently removed from the total RNA using a Ribo-Zero™ Magnetic Kit for Bacteria (Epicentre Biotechnologies, USA). The efficiency of this rRNA removal and the quality of the remaining RNA were measured again by BioAnalyzer 2100 Total RNA Nano Kit (Agilent Technologies Inc., USA). The remaining rRNA contamination that could be visualized was below 5%. RNA-seq libraries were prepared using the NEBNext® Ultra™ Directional RNA Library Prep Kit. Sequencing was carried out on an Illumina MiSeq platform at KAUST Bioscience Core Laboratory.

Quality control (QC) of the transcriptomic reads was performed in the following order. 1) Trimming reads with Trimmomatic (version 0.33) (Bolger *et al.*, 2014), 2) error correcting with SPAdes (version 3.0.0) (Nurk *et al.*, 2013), 3) remaining rRNA contamination removal with SortmeRNA (version

2.0) (Kopylova *et al.*, 2012). Only overlapping pair-end reads were considered. INDIGO-predicted genes of strain ERBA were used to map the qualified transcripts using Bowtie2 (version 2.2.5) (Langmead and Salzberg, 2012). After mapping, reads from the SAM file were converted to read counts. Differential expression analysis of RNA-seq data across three conditions was performed using the edgeR Bioconductor package (version 3.2) (Robinson *et al.*, 2010). An MDS plot was generated to visualize sample relations. Exact conditional tests were conducted to identify differential expression between different growth conditions. Significantly differentially expressed genes were defined as those with Exact Test *p value* <0.01. The 3rd quartile of all $\log_{10}$ FPKM values obtained is 2.541 (the actual FPKM value is 347.536 accordingly). Thus, genes with FPKM value greater than 350 are considered highly expressed genes.

2.3 Results and discussion

2.3.1 Cell morphology and characteristics

In 1987, the founding member of the extremely halophilic methanogens, *Methanohalobium (Mhb.) evestigatum,* was isolated from a microbial mat of Sivash Lake in Crimea. The cells of *Mhb. evestigatum* are irregular cocci or polygonal flattened with a diameter of 0.2 – 2μm (Zhilina and Zavarzin, 1987) (Figure 10A). Ever since, no new cultures affiliated with this genus have been described.

28 years later, for this thesis, we isolated a new methanogenic archaeon, designated as strain ERBA, from the brine-seawater interface of Erba Deep in

the Red Sea. This new methanogen shares 98% 16S ribosomal rRNA similarity with *Mhb. evestigatum* suggesting a close relationship with this species. The cells of strain ERBA showed distinct phenotypic characteristics – they are cocci with diameters of 0.5 – 5μm that form cell clusters and buddings as shown in Figure 10B and C. Motility was not observed for either *Methanohalobium* strains.

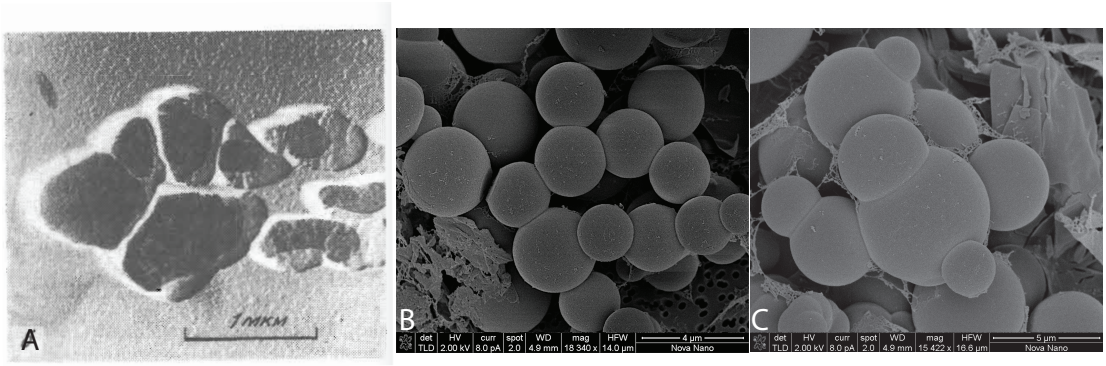


Figure 10. Scanning electron micrograph of cells of *Mhb. evestigatum* (A) (Zhilina and Zavarzin, 1987) and *Mhb.* strain ERBA. (Note formation of cell clusters and buddings in B and C, respectively.)

Both strains share the same substrates (methanol and mono-, di- and tri-methylamines) for growth and methanogenesis. B-vitamins were found to stimulate their growth. They also share a similar suitable growth pH range. However, they differ considerably in growth temperatures and salinities. The cells of *Mhb. evestigatum* grow optimally at 25% NaCl under 50°C and do not grow below 15% NaCl. The optimal growth salinity and pH of the strain ERBA are very similar to the parameters of the environment in which this strain was isolated. It optimally grows at around 37°C and with 10% NaCl. Under laboratory cultivating conditions and without yeast extract and peptone, strain ERBA showed a much wider range of NaCl adaptability (4% - 29%) relative to

Mhb. evestigatum. In summary, *Mhb. evestigatum* is a halophilic thermophile. Strain ERBA is a halophilic mesophile (refer also to Table 6). Due to the overall scarcity of *Methanohalobium* isolates and our ability to establish cultures of these two pioneer members of halophilic methanogens, they can now serve as excellent model organisms to understand the unique biology of halophilic methanogens that is necessary for adapting to extreme environments.

Table 6. Physiological characteristics of *Methanohalobium* strains.

Characteristic	*Methanohalobium evestigatum*	*Methanohalobium* strain ERBA
Isolation	Zhilina & Zavarzin (1987)	This study
Cell type	Polygonal, irregular cocci	Cocci
Size (μm)	0.2-2 (single), 5-10 (aggregate)	0.5 – 5
Temp range (°C)	20 - 60	20 - 45
Optimal temp (°C)	~ 50	33-37
pH range	6.0 - 8.0	6.0 - 8.0
Optimal pH	7.0 - 7.5	7.0 - 7.8
NaCl range (%)	15 - 30	4 – 29
Optimal NaCl (%)	25	~10
Substrates	Methanol	Methanol
	MMA DMA TMA*	MMA DMA TMA
Growth factors	B-vitamins	Stimulate growth: B-vitamins, yeast extract, peptone
Source	Sivash Lake, microbial mat	Erba Deep (brine-seawater interface), Red Sea
Source salinity	n.d.	8-10.6%

*MMA, DMA and TMA stand for monomethylamine, dimethylamine trimethylamine, respectively.

2.3.2 General genome characteristics

The genome sequence of *Mhb. evestigatum* was sequenced and released in 2010 by the DOE Joint Genome Institute, but there is no associated genome description publication to date. It consists of a circular chromosome (NC_014253.1) with a size of 2.24Mb and a 0.16Mb plasmid (NC_014254.1) (both depicted in Figure 11). According to GenBank and the respective

Integrated Microbial Genomes server (IMG) record, the chromosome contains approximately 2318 predicted genes (under the locus tag Metev) and the plasmid bears 119 predicted genes.

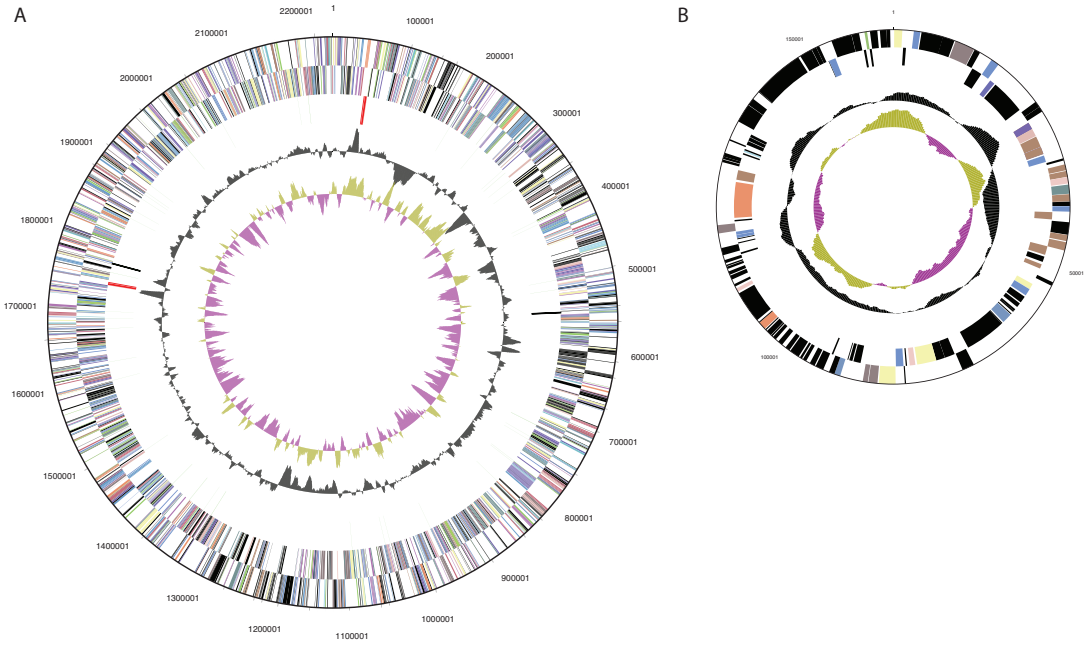


Figure 11. Circular representation of the *Methanohalobium evestigatum* DSM3721 genome. A) The chromosome; B) The plasmid. The circles show (outermost to innermost): DNA coordinates (black), genes on forward strand (color coded by COG categories), genes on reverse strand (color coded by COG categories), RNA genes (including tRNAs in green, rRNAs in red and other RNAs in black), GC content, and GC skew.

The high-quality draft genome of strain ERBA is 2.04 Mb long and consists of 26 scaffolds. It contains approximately 2159 predicted genes according to INDIGO annotation (under the locus tag ERBA; 2080 predicted genes according to IMG prediction). Based on the presence of all 104 conserved archaeal single-copy genes identified by AmphoraNET, we estimate that the

draft genome of ERBA is complete. There is no evidence for the presence of a plasmid in this version of ERBA draft genome.

Table 7 provides an overview of both *Methanohalobium* genomes. They share overall similar statistical features. The genome of *Mhb. evestigatum* has two 16S-23S-5S rRNA operons. However, the draft genome of strain ERBA only has one complete 16S-23S-5S rRNA operon and one additional 5S rRNA at one end of contig 14, which suggests a truncated version of the rRNA operon and could be caused by the incompleteness of the draft genome. The G+C contents of these two *Methanohalobium* species are similar. These two are among genomes of the lowest G+C contents within the methanogenic order of *Methanosarcinales*.

Table 7. Overview of *Methanohalobium* genomes. (The total number of genes and predictions were based on the Integrated Microbial Genomes (IMG) server annotation, thus they differ slightly from INDIGO annotation.)

	Methanohalobium evestigatum		*Methanohalobium* strain ERBA (draft)	
	Number	% of Total	Number	% of Total
DNA, total number of bases	2406232	100.00%	2044989	100.00%
DNA coding number of bases	2114370	87.87%	1800223	88.03%
DNA G+C number of bases	876469	36.42%	767031	37.51%
Plasmid Count	1		0	
CRISPR array	0		2 (86 spacers)	
Genes total number	2437	100.00%	2132	100.00%
Protein coding genes	2378	97.58%	2080	97.56%
Pseudo Genes	124	5.09%		
RNA genes	59	2.42%	52	2.44%
rRNA genes	6	0.25%	4	0.19%
5S rRNA	2	0.08%	2	0.09%
16S rRNA	2	0.08%	1	0.05%
23S rRNA	2	0.08%	1	0.05%
tRNA genes	51	2.09%	45	2.11%
Other RNA genes	2	0.08%	3	0.14%
Protein coding genes with function prediction	1540	63.19%	1690	79.27%

without function prediction	838	34.39%	390	18.29%
Protein coding genes connected to KEGG pathways	665	27.29%	656	30.77%
not connected to KEGG pathways	1713	70.29%	1424	66.79%
Fused Protein coding genes	144	5.91%	56	2.63%
Protein coding genes coding signal peptides	80	3.28%	62	2.91%
Protein coding genes coding transmembrane proteins	116	4.76%	451	21.15%

The genome of *Methanohalobium evestigatum* Z-7303 encodes 51 tRNAs for all 20 amino acids plus, additionally, pyrrolysine, and two other non-coding RNAs (one ffs RNA -4.5S, RNA component of the signal recognition particle (SRP) ribonucleoprotein complex; one rnpB ncRNA-RNase P RNA, a catalytic subunit of RNAse P). *Methanohalobium sp.* strain ERBA has 47 tRNAs (including tRNA-trp and tRNA-pyl by manual identification, of which 45 tRNAs could be identified by INDIGO and IMG) to decode 21 amino acids. By using the two non-coding RNAs in *Mhb. evestigatum* as query, their counterparts in the strain ERBA could also be identified.

These two genomes have the corresponding tRNA synthetase for each tRNA. Like other typical methylated compound utilizing methanogens, both of these *Methanohalobium* strains possess corresponding $tRNA_{pyl}$ synthetase and all enzymes necessary for the biosynthesis of pyrrolysine (Gaston, Zhang, *et al.*, 2011). In contrast to other members in the order of *Methanosarcinales* which have their *pylTSBCD* genes form one cluster (Gaston, Jiang, *et al.*, 2011), uninterrupted by other genes, *Methanohalobium* species's *pylTS* genes and *pylBCD* genes are separated at about 0.6Mb distance in the case of *Mhb. evestigatum* and located on the middle of different scaffolds in the draft genome of strain ERBA.

Although the tRNA gene for selenocysteine is not present, both *Methanohalobium* strains have a truncated version of the selenocysteine-specific elongation factor (selB)/ elongation factor Tu. This type of elongation factor is both a 3' and 5' ends truncated version of *selB*, which is widespread in non-selenocysteine utilizing methanogens.

The two different *Methanohalobium* strains share 97% tetranucleotide identity, a relatively low sequence identity (81% at nucleotide level, and 84% at the amino acid level) over the overlapped genome regions. Based on an ANI threshold of less than 95%, that corresponds to a 70% DNA-DNA hybridization for identifying novel species (Goris *et al.*, 2007), strain ERBA therefore represents a novel species within genus *Methanohalobium*. A large number of CDS are orthologs and syntenic, showing a relatively high degree of functional conservation in these two genomes (see also Table 8). The two genomes share a substantial fraction of their genetic repertoire with a core genome of 1743 genes. With respect to species-specific genes, 546 predicted genes are unique to *Mhb. evestigatum* and 364 are unique to strain ERBA. 35.7% (i.e., 130) of the total unique genes predicted in ERBA and 33.2% (i.e., 181) of the total unique genes predicted in *Mhb. evestigatum* encode hypothetical proteins. 7.4% (27 genes in ERBA) and 10.8% (59 genes in *Mhb. evestigatum*) are predicted genes for transposases. The majority of unique genes in the *Mhb. evestigatum* genome are concentrated on 7 genomic islands ranging from 8 Kb to 20 Kb in length. In contrast, there is only one identifiable genomic island in the ERBA genome of 43 Kb in length (from ERBA_2244 to ERBA_2295) carrying a large number of hypothetical protein coding genes.

Table 8. Gene order conservation (synteny) of *Methanohalobium* strain ERBA compared to *Methanohalobium evestigatum*.

Subject	Query	ANI (%)	AAI (%)	Number of CDS	
				Orthologs	Syntenic
Mhb. evestigatum	Strain ERBA	81.13	84.35	1724	1642

ANI: Average nucleotide identity based on the BLAST algorithm.

AAI: Average amino acid identity.

These two strains have very similar gene COG functional category distributions (as listed in Table 9). The top represented COG categories (shaded in grey in Table 9) are General function prediction only [R] and Function unknown [S], Translation, ribosomal structure and biogenesis [J], Replication, recombination and repair [L], Energy production and conversion [C], Coenzyme transport and metabolism [H], Amino acid transport and metabolism [E] and Inorganic ion transport and metabolism [P]. Notably, *Methanohalobium* species contain the most transposases in methanogens known to date. Transposase mediates cut-and-paste transposition of DNA sequence in a very efficient manner (Lampe *et al.*, 1996; Montaño and Rice, 2011). It implies an important role of transposition and gene recombination in shaping *Methanohalobium* genomes under hypersaline conditions.

Table 9. Number of genes associated with the general COG functional categories.

Description and COG class	**COG counts [a] and percentage of protein-coding genes**					
	***Mhb. evestigatum* chr**		***Mhb. evestigatum* chr+plsm**		**Strain ERBA**	
	value	% of total	value	% of total	value	% of total
RNA processing and modification [A]	2	0.11	2	0.11	2	0.11
Chromatin structure and dynamics [B]	4	0.22	4	0.22	4	0.23
Energy production and conversion [C]	151	8.41	151	8.16	157	8.91
Cell cycle control, cell division, chromosome partitioning [D]	14	0.78	16	0.86	14	0.79
Amino acid transport and metabolism [E]	133	7.41	133	7.19	138	7.83
Nucleotide transport and metabolism [F]	49	2.73	49	2.65	51	2.89

Carbohydrate transport and metabolism [G]	60	3.34	62	3.35	61	3.46
Coenzyme transport and metabolism [H]	132	7.35	133	7.19	146	8.28
Lipid transport and metabolism [I]	22	1.22	23	1.24	22	1.25
Translation, ribosomal structure and biogenesis [J]	159	8.85	162	8.76	159	9.02
Transcription [K]	88	4.9	91	4.92	81	4.59
Replication, recombination and repair [L]	150	8.35	162	8.76	97	5.5
Cell wall/membrane/envelope biogenesis [M]	41	2.28	52	2.81	39	2.21
Cell motility [N]	25	1.39	26	1.41	23	1.3
Posttranslational modification, protein turnover, chaperones [O]	82	4.57	85	4.59	83	4.71
Inorganic ion transport and metabolism [P]	80	4.45	82	4.43	97	5.5
Secondary metabolites biosynthesis, transport and catabolism [Q]	15	0.84	15	0.81	12	0.68
General function prediction only [R]	239	13.31	240	12.97	243	13.78
Function unknown [S]	202	11.25	205	11.08	204	11.57
Signal transduction mechanisms [T]	97	5.4	103	5.57	94	5.33
Intracellular trafficking, secretion, and vesicular transport [U]	23	1.28	24	1.3	25	1.42
Defense mechanisms [V]	28	1.56	30	1.62	11	0.62
total gene numbers in COGs	1796		1850		1763	

[a]Some genes belong to more than one functional category.

The main translation initiation codon used by *Methanohalobium* is AUG (accounting for about 85% of all predicted protein-coding genes located on chromosomes). Three copies of tRNA-methionine are identified in each genome. GUG and UUG are also found in a lower amount as putative translation start codons. Ochre (UAA) and opal (UGA) codons are used as major stop codons. Ochre codons are used in higher frequencies (63% - 68%) than opal codon (31% – 36%). Similar to other methylotrophic methanogens, the amber codon (UAG) is significantly underrepresented in *Methanohalobium* (less than 2%).

2.3.3 Phylogenetic placement of the Methanohalobium

Phylogenomic analysis revealed that the two *Methanohalobium* species clustered together and formed a sister group with *Methanosalsum zhilinae*, an alkaliphilic, halophilic and methylotrophic methanogen isolated from the Wadi el Natrun in Egypt (Mathrani *et al.*, 1988) (Figure 12). This lineage is

separated from the majority of other methylotrophic methanogens: moderately halophilic methanogens *Methanohalophilus* and the slightly halophilic methanogens *Methanococcoides*, *Methanolobus* and *Methanomethylovorans*. All these branches received strong bootstrap support (>96%).

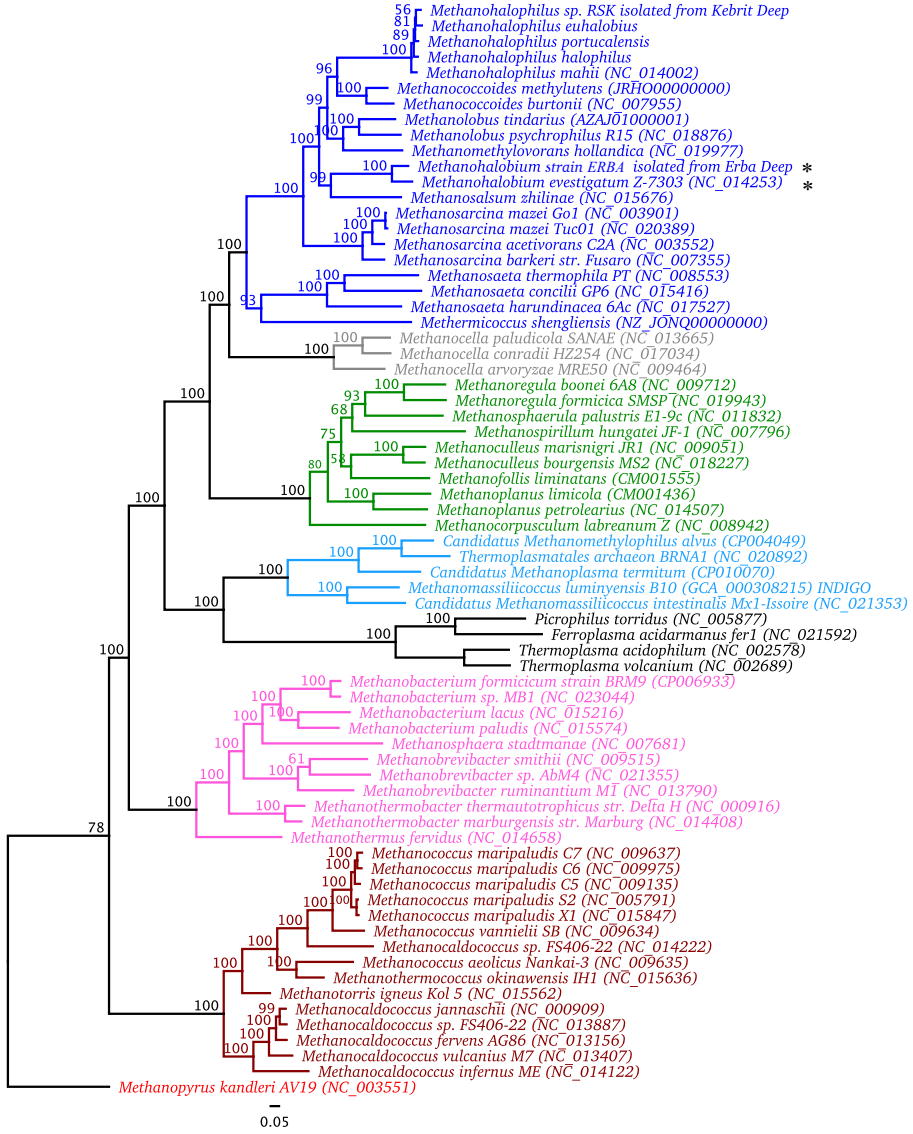


Figure 12. Phylogenomic inference of *Methanohalobium* in association with other methanogens. The tree is rooted to *Methanopyrus kandleri* (in red). Asterisks indicate *Methanohalobium*. *Methanosarcinales* are shown in blue, *Methanocellales* are shown in grey, *Methanomicrobiales* are shown in green, *Methanomassiliicoccales* are shown in light blue, *Methanobacteriales* are shown in purple, *Methanococcales* are shown in ochre.

2.3.4 Origin of replication

The *Mhb. evestigatum* genome possesses three copies of the *orc1/cdc6* (origin recognition complex/cell division cycle 6) gene, while the draft genome of strain ERBA has two copies (ERBA_02049 and ERBA_02298), which are homologs of Metev_0001 and Metev_2049 in *Mhb. evestigatum*, respectively. Gene loci Metev_0001 and ERBA_02049 encodes the conserved Orc1/Cdc6.1 that is suggested to be the main initiator protein (Borrel *et al.*, 2014). The three *orc1/cdc6* genes in *Mhb evestigatum* are located on the same strand. However, the two *orc1/cdc6* genes in strain ERBA are located on a different strand. Unique to ERBA, a gene for plasmid pRiA4b ORF-3-like protein (ERBA_01294) is present. This protein has closest BLASTp hits from uncultured ANME-1 archaeon (CBH38831, 58%) and 'Lokiarchaeum' (KKK45145, 56% identity), which could potentially play a role in DNA replication in benthic microorganisms.

Chromosomal origin (oriC) regions of each genome were predicted using Ori-Finder2 (Luo *et al.*, 2014). *Methanohalobium* evestigatum exhibits potentially multiple origins of replication according to this prediction. Three to four Origin Recognition Box (ORB) motifs could be found in the neighborhood of only one *orc1/cdc6* gene in each genome (Metev_0001 in *Mhb. evestigatum* and ERBA_02049 in strain ERBA) (Table 10).

Table 10 OriC and ORBs motifs in the *Methanohalobium* genomes.

oriC location (nt)	**ORB Sequence**	**Strand**	**Length**	**Comment**	**Distance to *orc1/cdc6.1* (nt)**	**Position**
Mhb. evestigatum (2242116 - 357)	**TCCA**ATT**GAAA**TAAT**GGGGT**	-	20 bp	upstream *orc1/cdc6.1*	448	2,242,207 - 2,242,226
Size: 559 nt. GC level: 24.87%.	**TCCAGT**A**GAAA**TAAT**GGGG**A	+	20 bp	upstream *orc1/cdc6.1*	254	84 - 103
	C**CCA**CTT**GAAA**CAAA**GGGGT**	+	20 bp	upstream *orc1/cdc6.1*	173	165 - 184
	TCCATT**GGAA**TTAAAGCAGT	+	20 bp	upstream *orc1/cdc6.1*	124	214 - 233
Mhb. evestigatum (48221 - 54207)	**TCCA**TT**GGAA**TTCTG**ACGGT**	-	20 bp		47632	49227 - 49246
Size: 5987 nt. GC level: 51.01%.	**TCCA**ACGG**GAA**G**AACGGCCT**	+	20 bp		50829	52424 - 52443
	TTAACCTG**ATA**T**CAAGGGTT**	–	20 bp		52386	53981 - 54000
Mhb. evestigatum (245524 - 248218)	**TACAG**A**GGAAATA**CAATCG**A**	+	20 bp		205041	246636 - 246655
Size: 2695 nt. GC level: 32.8%.	TATA**AGTGAAACAA**GA**GGGT**	+	20 bp		245157	246752 - 246771
	TCTACTTGAAACCAAGTTTT	+	20 bp		246413	248009 - 248028
Mhb. Strain ERBA (contig2 26842-27483)	TGCAGTTGAAATAA**AGGGGT**	-	20 bp	upstream *orc1/cdc6.1*	178	27,020 - 27,039
Size: 642 nt. GC level: 26.01%.	**TCCAGTAGAAA**CAATG**GGGGT**	-	20 bp	upstream *orc1/cdc6.1*	257	27,099 - 27,118
	TCCAGTAGAAACAATG**GGGGT**	+	20 bp	upstream *orc1/cdc6.1*	452	27,294 - 27,313
Archaea consensus ORB	C**TTCCAGTGGAAA**CG**AAGGGGGT**					
Methanomicrobia consensus ORB	**TCCA**GTG**GAAA**CA**AAGGGGT**					

2.3.5 Amber codon usage and putative Pyl-containing proteins

Methanohalobium could translate the amber stop codon (UAG) as pyrrolysine (Pyl) (Gaston, Jiang, *et al.*, 2011). In searching for Pyl-containing proteins in *Methanohalobium* genomes, all the INDIGO-annotated genes with TAG as stop codon were selected (31 genes in strain ERBA and 46 genes in *Mhb. evestigatum*). Their existing translation frames were extended to their 3' flanking regions up to the next non-TAG stop codon. The new CDS will be predicted as potential Pyl-incorporating proteins only if it shares significant BLASTp hits with proteins in NCBI Non-redundant RefSeq protein records. According to this method, 10 putative Pyl-incorporating proteins are predicted for strain ERBA (accounting for 32% predicted genes with a TAG stop codon in total) and 12 are predicted for *Mhb. evestigatum*, amounting to a total of 26% predicted genes with a TAG stop codon (see also Table 11). Consistent with previous finding on methylotrophic methanogens (Borrel *et al.*, 2014; Gaston, Jiang, *et al.*, 2011), *Methanohalobium* could incorporate Pyl to both substrate-specific methylamine methyltransferases in the methanogenesis pathway and to enzymes that are not related to this pathway.

Table 11. putative Pyl-containing proteins in Methanohalobium.

In strain ERBA

Locus ID	Annotation	Extended 3' region
ERBA_01734	Trimethylamine methyltransferase MttB	ERBA_01733
ERBA_01669	Trimethylamine methyltransferase MttB	ERBA_01668
ERBA_00961	Dimethylamine methyltransferase MtbB2	5' incomplete, and ERBA_00960
ERBA_01991	Monomethylamine methyltransferase MtmB	ERBA_01992
ERBA_01717	Phage shock protein C (PspC)	46 AA until TGA
ERBA_01435	Phospholipiddiacylglycerol acyltransferase	22 AA until TAA
ERBA_01441	Hypothetical protein	2 Pyls, 77 aa until TAA
ERBA_00529	ArsR family transcriptional regulator	ERBA_00530
ERBA_01433	Aerobic respiration control sensor protein ArcB	ERBA_01434

ERBA_01589	4Fe-4S ferredoxin iron-sulfur binding domain containing protein	Part of ERBA_01588 (not in-frame)

In *Mhb. Evestigatum*

INDIGO Locus ID	**Annotation**	**Extended 3' region**
InMeves_02220	Dimethylamine methyltransferase MtbB1	InMeves_02219
InMeves_01613	Dimethylamine methyltransferase MtbB1	InMeves_01612
InMeves_01195	Monomethylamine methyltransferase MtmB	InMeves_01194
InMeves_01504	Monomethylamine methyltransferase MtmB	InMeves_01503
InMeves_02227	Trimethylamine methyltransferase MttB protein	InMeves_02226
InMeves_00663	methanosis marker protein 12	InMeves_00662
InMeves_01443	HNH endonuclease	96 AA until TAA
InMeves_02433	3-ketosteroid-9-alpha-hydroxylase reductase subunit	57 AA until TAA
InMeves_01450	Uracil phosphoribosyltransferase protein	20AAa until TAA
InMeves_01108	DNA helicase	InMeves_01109
InMeves_02522	ABC transporter protein	InMeves_02521, cobalt transporter ATP-binding subunit
InMeves_01809	ArsR family transcriptional regulator	InMeves_01810

2.3.6 Coenzyme biosynthesis

Coenzyme M (2-mercaptoethanesulfonic acid) is the terminal methyl carrier and an essential coenzyme in methanogenesis (Taylor and Wolfe, 1974). Consistent with features of the proposed coenzyme M synthesis pathway in other Methanosarcinales (David E Graham *et al.*, 2009), *Methanohalobium* genomes do not have *comABC* genes but have homologues of the *Methanosarcina acetivorans* MA MA3297 gene (for L-cysteate synthesis), aspartate aminotransferase and sulfopyruvate decarboxylase to yield coenzyme M. In *Methanohalobium*, this gene appears to be a fused version, unlike the majority of methanogens including obligately methylotrophic methanogens (*Methanohalophilus*, *Methanolobus*, *Methanococcoides*, *Methanosphaera*, and *Methanomassiliicoccus*) that encode the ComD and ComE proteins by adjacent two individual genes. For coenzyme F_{420} biosynthesis, a complete set of *cof* genes (*cofC/D/E/G/H*) in this pathway was identified. The presence of 2-

isopropylmalate synthase, 3-isopropylmalate dehydratase (small and large subunits) and 3-isopropylmalate dehydrogenase genes suggest a pathway for coenzyme B biosynthesis.

Figure 13 presents an overview of the selected metabolic pathways, energy production and transport in *Methanohalobium*. These metabolic characteristics will be analyzed and compared in some of the following sections of the chapter.

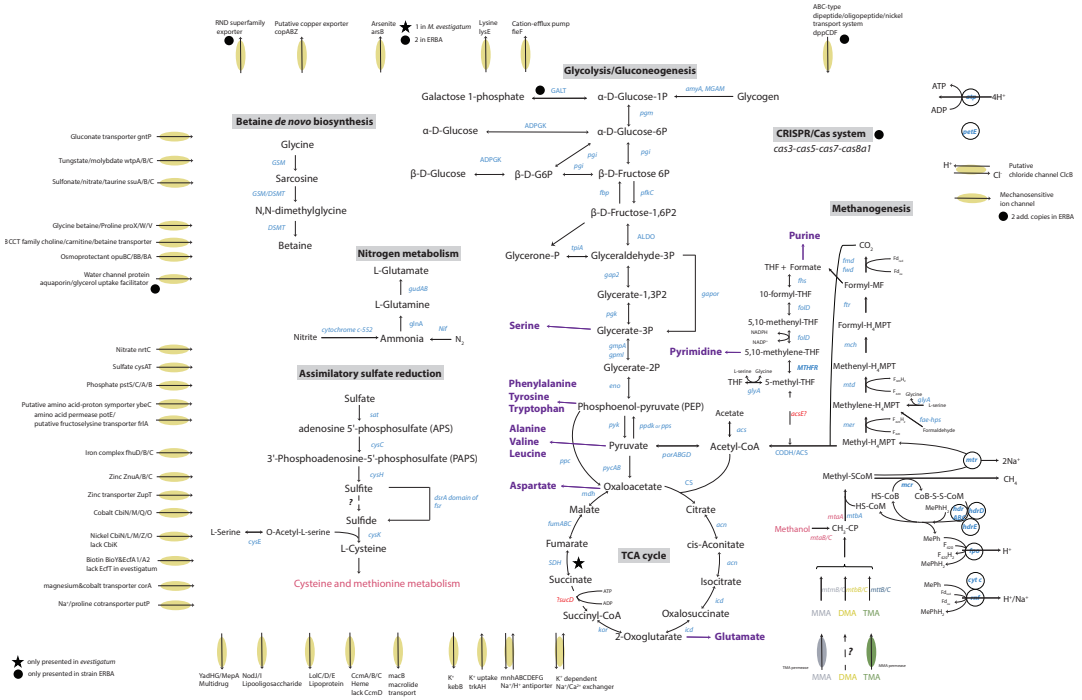


Figure 13. Metabolic reconstruction of selected pathways in *Methanohalobium*. Enzymes are shown in light blue and labeled with abbreviations. Membrane associated proteins and transporters are anchored in the membrane (indicated by the grey frame) with arrows to indicate the flow direction. Dashed lines and question marks represent pathways or steps for which no enzymes have been identified. Black circles indicate enzymes encoded by unique genes in ERBA. Stars indicate that the enzymes only present in *Mhb. evestigatum*.

2.3.7 Methanogenesis

Both *Mhb. evestigatum and Mhb.* strain ERBA are obligately methylotrophic methanogens. Their genomes provide the confirmation of their capacity to

grow on various methylamines and methanol. Both genomes encode trimethylamine (TMA) and monomethylamine (MMA) permeases that directly uptake these key substrates for methanogenesis. The absence of dimethylamine (DMA) permease indicates that their genome might encode unidentified dimethylamine permeases or that they might not directly uptake this methanogenesis substrate from the environment, but instead using the DMA from demethylation of TMA. Both *Mhb. evestigatum* and strain ERBA can grow on DMA under laboratory conditions, which support the former hypothesis. Both have the complete set of methanol, mono-, di- and tri-methylamine methyltransferases, corrinoid proteins and methylcobamideCoM methyltransferases (uroporphyrinogen decarboxylase) that enable the methyl group to transfer from the substrates to coenzyme M (CoM). Both genomes bear multiple copies of the majority of these genes. For example, when compared with *Mhb. evestigatum*, strain ERBA has several additional and unique copies of substrate-specific methyltransferases, corrinoid proteins and corrinoid protein:coenzyme M methyltransferase. This could provide strain ERBA with increased capacity and flexibility to utilize methane metabolic substrates in the polyextreme deep-sea brines.

Like other previously reported methylotrophic methanogens that grow on methylamines (Gaston, Jiang, *et al.*, 2011), all the mono, di- and tri-methylamine methyltransferases involved in methyl transfer during methanogenesis have an in-frame UAG codon and incorporate pyrrolysine (see also the section on Amber codon usage and putative Pyl-containing proteins in this chapter). However, in both genomes, a non-pyrrolysine member of

monomethylamine methyltransferase by annotation (Metev_1467 and ERBA_01108 with 83% amino acid sequence similarity) was observed with both lacking an internal amber TAG codon required for pyrrolysine incorporation. A similar case has been reported for a non-functional trimethylamine methyltransferase gene in *Methanococcoides burtonii* (Campanaro *et al.*, 2011). Based on low sequence similarities to other Pyl-incorporating monomethylamine methyltransferases and based on the fact that they are not associated with any gene clusters involved in the methyl-reduction pathway, this gene is unlikely to be a *bona fide* monomethylamine methyltransferase. The functionality and substrate specificity of this methyltransferase requires experimental verification.

All genes required for the reverse CO_2 reductive pathway have been identified in both genomes. Similar to many methylotrophic methanogens (Allen *et al.*, 2009; Spring *et al.*, 2010), *Methanohalobium* can oxidate methyl CoM through the reversal of the carbon dioxide reduction pathway. Like other H_2-independent methylotrophic methanogens, *Methanohalobium* appear to use CoB-CoM heterodisulfide reductase (Hdr) and $F_{420}H_2$ dehydrogenase (Fpo) for CoM regeneration. Genes for cytochrome *c* (cyt *c*) and the Rnf (***R**hodobacter* **n**itrogen **f**ixation) complex are also found in the genomes of *Methanohalobium*. This suggests that a similar energy-conserving electron transport pathway involving methanophenazine (MePh), ferredoxin (Fd) and cytochrome *c* to non-H_2-utilizing *Methanosarcina acetivorans* (Wang *et al.*, 2011) could also present in *Methanohalobium*. Plastocyanin (*petE*) genes are also present in the

Methanohalobium and may take part in the electron transfer (Walker *et al.*, 2010).

2.3.8 Carbohydrate metabolism

Acetyl-coenzyme A (acetyl-CoA) is a critical molecule in microbial metabolisms. In *Methanohalobium*, acetyl-CoA could be provided from various pathways, such as tetrahydromethanopterin-mediated reversal of CO_2 reduction pathway in methanogenesis or folate-mediated one-carbon (C1) metabolism via CODH/acetyl-CoA synthase complex, from acetate assimilation via acetyl-CoA synthetase or from decarboxylation of pyruvate catalyzed by pyruvate ferredoxin oxidoreductase. Thus, both methyl-compounds for methanogenesis and acetate could serve as carbon sources for *Methanohalobium*. Complete pathways for glycolysis and gluconeogenesis are also present in both.

Unlike the majority of methanogens that have a largely incomplete tricarboxylic acid cycle (TCA cycle), we predict that the extreme halophilic *Methanohalobium* could have the coding potential for a nearly complete TCA cycle. There are two pathways for biosynthesis of oxaloacetate: from acetyl-CoA via pyruvate-by-pyruvate synthase and pyruvate carboxylase, or from phosphoenolpyruvate-by-phosphoenolpyruvate carboxylase. Through the oxidative direction of the TCA cycle, citrate synthase, aconitase, and isocitrate dehydrogenase provide a pathway for synthesis of 2-oxoglutarate from oxaloacetate and acetyl-CoA. Continuing with the oxidative direction, succinyl-CoA could be synthesized by the 2-oxoglutarate ferredoxin oxidoreductase. In the reductive direction of the TCA cycle, both genomes contain malate

dehydrogenase and fumarate hydratase that enable the conversion of oxaloacetate to first malate and then fumarate.

ERBA seems to have lost the ability to convert fumarate to succinate. The succinate dehydrogenase subunit A and B coding genes are only present in *Mhb. evestigatum* (Metev_1450 and Metev_1449). However, ERBA has a homologue (ERBA_00011) of gene Mmah_1949 which has been speculated to encode fumarate reductase (succinate dehydrogenase) in *Methanohalophilus mahii* (Spring *et al.*, 2010) that might enable ERBA to complete this gap.

Whether *Methanohalobium* can achieve the interconversion of succinate and succinyl-CoA is questionable. A gene annotated as succinyl-CoA ligase ADP-forming subunit alpha in the genome of strain ERBA (ERBA_00934, with an orthologue Metev_1921 in *Mhb. evestigatum* annotated as acetyl coenzyme A synthetase) could serve as a candidate for this conversion. Both genes encode proteins containing a succinyl-CoA ligase-like flavodoxin domain and a CoA binding domain (same as the succinyl-CoA synthetase alpha subunit).

In one-carbon metabolism, phosphoribosylglycinamide formyltransferase (Fhs), bifunctional 5,10-methylene-tetrahydrofolate dehydrogenase (FolD), and 5,10-methylenetetrahydrofolate reductase (MTHFR) provide a pathway to generate 5-methyltetrahydrofolate (5-methyl-THF) from formate (or reverse) in *Methanohalobium*. However, 5-methyltetrahydrofolate:corrinoid/iron-sulfur protein methyltransferase (AcsE, EC 2.1.1.258) could not be identified, thus it is not clear which corrinoid methyl transferase in *Methanohalobium* could link 5-methyl-THF to acetyl-CoA. Alternatively, similar to the way in which

methyltetrahydromethanopterin (CH_3-H_4MPT) serves as methyl donor to acetyl-CoA, we postulate that 5-methyl-THF might directly act with the CODH/ACS complex for acetyl-CoA formation.

Methanohalobium might be able to store carbon as glycogen (starch) and utilize it for glycogenolysis. Both genomes have classic starch degradation genes including alpha-amylase (*amyA*), glycogen-debranching protein (*AGL*), glycoamylase (*MGAM*), and glycogen phosphorylase (alpha-glucan phosphorylase). So far, this *Methanohalobium*-type glycogen phosphorylase (Metev_0787) has only been found in *Methanosarcina* species and *Methanolobus tindarius* within methanogens. Glycogen synthase (ERBA_01248) and galactose-1-phosphate uridylyltransferase (ERBA_01247) are only present in the genome of strain ERBA. It indicates that only strain ERBA could perform *de novo* synthesis of glycogen and convert between galactose 1-phosphate and D-Glucose-1P.

2.3.9 Sulfur metabolism

Methanohalobium species have a rather complex sulfur network in comparison to many other methanogens. Genes for sulfate ABC transporters are present in both genomes. In their assimilatory sulfate reduction pathway, sulfate adenylyltransferase, adenylylsulfate kinase, phosphoadenosine phosphosulfate reductase could reduce sulfate to sulfite through intermediates adenylyl sulfate (APS) and 3'-phosphoadenylyl sulfate (PAPS). Both genomes carry multiple copies of phosphoadenosine phosphosulfate reductase gene (5 copies in *Mhb. evestigatum* and 3 copies in ERBA). *Methanohalobium* could also possibly convert thiosulfate to sulfite. Genes encoding rhodanese domain-containing

protein (ERBA_00849) were identified in both genomes. This gene is unique to *Methanohalobium* in the archaeal domain. All its existing homologs are from the bacteria domain with a low BLASTp similarity of less than 36% in the NCBI-nr database. This gene is adjacent to a gene annotated as hydroxyacylglutathione hydrolase (ERBA_00850), which contains two rhodanese domains.

Although the dissimilatory sulfate reduction or using sulfite as sole sulfur source has not been experimentally characterized and confirmed in *Methanohalobium*, both species possess the coenzyme F_{420}-dependent sulfite reductase (Fsr) orthologues that have previously been found only in 3 strictly hydrogenotrophic thermophilic methanogens (*Methanocaldococcus jannaschii, Methanopyrus kandleri* and *Methanothermobacter thermautotrophicus*) as well as the Antarctic archaeon *Methanococcoides burtonii*. It has been previously characterized only in *Methanocaldococcus jannaschii* (Johnson and Mukhopadhyay, 2005; 2008). Fsr contains a H_2F_{420} dehydrogenase N-terminal half and a dissimilatory-type siroheme sulfite reductase C-terminal half. The expression of this enzyme in *Methanocaldococcus jannaschii* is induced by sulfite (Johnson and Mukhopadhyay, 2005). In strain ERBA, gene ERBA_00438 had very low expression (FPKM<86) in the media without supplemental sulfite. Methanogenic archaea are usually sensitive to sulfite (Balderston and Payne, 1976). We hypothesize that this enzyme could be used in sulfite detoxification and/or to convert sulfite to sulfide in *Methanohalobium*. With the advantage of numerous recent published methanogen genomes, we looked into NCBI-nr protein database and identified

an additional 36 Fsr homologs. Taken together, Fsr homologs are more widespread in hydrogenotrophic methanogens than in methylotrophic methanogens (which are only found in *Methanococcoides*-2 species, *Methanohalobium*-2 species, *Methanohalophilus*-1 species and *Methermicoccus*-1 species) and in a putative methane-oxidizing archaeon GZfos27A8 (*Methanohalobium evestigatum* and *Methanococcoides methylutens* each has two copies of Fsr gene). Furthermore, ERBA also contains a DsrE/DsrF/DrsH-like domain-containing protein (ERBA_02346) that is potentially involved in intracellular sulfur reduction. Finally, sulfide could be converted to L-cysteine by cysteine synthase A (Metev_1879) for protein synthesis.

In addition, genes encoding FAD-dependent pyridine nucleotide-disulfide oxidoreductase and sulfite oxidase-like oxidoreductase were also identified in both genomes. *Methanohalobium* might also be able to perform dimethyl sulfoxide (DMSO) reduction, since a putative anaerobic dimethyl sulfoxide reductase is encoded in both genomes.

Methanohalobium could depend on exogenous reduced sulfur for biosynthesis, since both species possess amino acid transporter orthologs. Both genomes encode amino acid permeases. A species-specific amino acid permease is also present in ERBA (ERBA_00806) that is most similar with its homologue in a haloarchaeon *Candidatus* Halobonum tyrrellensis (WP_023393102, 57% identity). A unique feature of strain ERBA is the potential ability to uptake peptides. Genes encoding for oligopeptide transport ATP-binding protein (OppF, ERBA_00611), dipeptide transport ATP-binding protein (DppD, ERBA_

00612), dipeptide transport system permease protein (DppC, ERBA_00613) are present.

2.3.10 Nitrogen metabolism

Members of methanogenic archaea are known to utilize ammonia as a nitrogen source or/and fixing dinitrogen (Belay *et al.*, 1984; Galagan *et al.*, 2002). Ammonia transporter genes could not be identified in either *Mhb. evestigatum* or strain ERBA. It is likely that the ammonia they acquire is passively diffused through the cell membrane. *Methanohalobium* could assimilate ammonia by glutamine synthetase, glutamate synthetase, and glutamate dehydrogenase, as both genomes encode these enzymes.

Although ERBA cultures can grow in the cultivation media without ammonia for several generations, the genomic evidence for *Methanohalobium* being able to fix dinitrogen is not very strong. Similar to the genome of *Methanococcoides burtonii* (Allen *et al.*, 2009), they have minimal gene components for dinitrogen fixation—Mo-nitrogenase iron protein subunits (*nifH* and *nifD*), FeS cluster assembly scaffold protein (*nifU*), FeMo cofactor biosynthesis protein (*nifB*), dinitrogenase iron-molybdenum cofactor biosynthesis (*nifx-nifB*), and nitrogen regulatory protein P-II (*glnB*) are present.

The presence of ABC-type nitrate/sulfonate/bicarbonate transport system genes in *Methanohalobium* suggests that they could uptake nitrate from their environment. *Methanohalobium* do not have identifiable urea ABC transporters.

The following proteins are present in both genomes and could potentially participate in the nitrogen metabolism or they could be involved in nitrosative stress protection in *Methanohalobium*: cytochrome c554 domain containing proteins (ERBA_02238, involves in the oxidation of ammonia), nitrite (sulfite) reductase (ERBA_00366, involves in the biosynthetic assimilation of sulfur and nitrogen), hydroxylamine reductase HCP (ERBA_00432, catalyzes the reduction of hydroxylamine to form NH_3), nitric oxide reductase activation protein (ERBA_00063), and nitric oxide reductases (ERBA_00551 and ERBA_02128).

2.3.11 Osmoadaptation

Osmotic stress is a common challenge to methanogens. Therefore, osmoregulation and adaptation is of essential for halophilic methanogens to thrive in hypersaline environments. Genome sequence analysis revealed the presence of a large and diverse group of protein coding genes that is potentially associated with the adaptation of *Methanohalobium* to high salinity.

Both genomes harbor genes for multiple copies of mechanosensitive ion channel proteins (ERBA having two additional copies) that could function as osmosensors and 'safety-valves' in *Methanohalobium* (Kloda and Martinac, 2001; Martinac, 2004). They also could use adenylate cyclase B (Kimura *et al.*, 2005) and multiple osmosensing transporters as osmosensors. The putative chloride channel protein ClcB (H^+/Cl^- antiporter) in *Methanohalobium* (ERBA_00936) hypothetically plays roles in maintaining cell volume and homeostasis, organic solute transport, etc.

The “salt-in” strategy (accumulating K^+ and C^- ions to prevent water efflux and expelling Na^+ ions) and the “compatible-solute” strategy are considered the two main methods to enable microorganisms to keep their cytoplasm at least isoosmotic with the extracellular environment (Oren, 1999) and to thus cope with osmotic stress. *Methanohalobium* species seem to use both of these two approaches to enhance their capability to adapt to hypersaline environments.

Under the **“salt-in” strategy**, *Methanohalobium* could import K^+ and exclude Na^+. For potassium intake, both genomes have multiple gene copies for low-affinity potassium transporters (TrkAH). ATP-driven K^+ import systems could not be identified. Predicted cation-transporting ATPase ERBA_01124 (with an orthologue in *Mhb. evestigatum*) and ERBA_01562 (unique to ERBA and with a pfam00689 domain) might be involved in Na^+/K^+, H^+/K^+, Ca^{2+} and Mg^{2+} transport in *Methanohalobium*. For sodium efflux, multi-subunit Mnh Na^+/H^+ antiporter genes are present in 7-gene clusters (from ERBA_00856 to ERBA_00862). Both genomes appear to have multiple copies of mechanosensitive ion channel genes (3 copies in *Mhb. evestigatum* and 5 copies in ERBA). The protein encoded by this gene has been suggested to play a role in osmotic shock by non-specific ion loss (Levina *et al.*, 1999).

Under the **“compatible-solute” strategy** on the other hand, both *Methanohalobium* species have the genomic capacity to uptake compatible solutes. Glycine betaine/proline transporters (ProX, ProW and ProV), ABC-type proline/glycine betaine transport systems (OpuBB and OpuBA) and BCCT family choline/carnitine/betaine transporter are present in both genomes.

Glycine betaine is a very efficient osmolyte found in all three domains of life. It is generally accumulated at high cytoplasmic concentrations and acts as an osmoprotectant to osmotic balance of living cells (J E Graham and Wilkinson, 1992). A gene for water channel protein aquaporin/glycerol uptake facilitator is only present in strain ERBA (ERBA_01336). This protein could be related to osmoadaptation as it could transport water or glycerol to help cells balance osmotic conditions.

In addition to taking up osmolytes directly from the environment, they could also be synthesized either by oxidation reactions from choline or by methylation from glycine (Nyyssola *et al.*, 2000). From a bioenergetic perspective, synthesizing compatible solutes, such as glycine betaine, are more costly for cells than directly updating osmolytes or inorganic compounds to help organisms cope with osmotic stress (Oren, 1999).

Despite this energetic cost, a pathway to *de novo* biosynthesis of osmoprotectant glycine betaine through methylation of glycine-by-glycine sarcosine methyltransferase (GSM) and sarcosine dimethylglycine methyltransferase (SDMT, also called malonyl-acyl-carrier protein O-methyltransferase)(Nyyssola *et al.*, 2000) is predicted to appear in both. These two key enzymes are located in the vicinity of a S-adenosylmethionine synthase, adenosine kinase (*adoK*) and adenosylhomocysteinase with predicted functions of providing S-adenosyl-L-methionine (AdoMet) as the methyl donor and of removing S-Adenosyl-L-homocysteine (AdoHcy), the demethylation product and the competitive inhibitor of methyltransferases

(Shu-Jung Lai and Mei-Chin Lai, 2011). This five-gene cluster is well conserved and has so far only been found in both the *Methanohalobium* species (loci in strain ERBA: ERBA_00872 – ERBA_00868) and all known moderately halophilic *Methanohalophilus* species as illustrated in Figure 14. The deduced protein sequences of GSM and SDMT genes in *Methanohalobium* and *Methanohalophilus* species share high similarities (>85%). These genes might be acquired by their common ancestors by horizontal gene transfer from bacteria. Constant high expression of this gene cluster was shown in the transcriptome dataset with a FPKM value greater than 800 under all tested salinity conditions (5%, 10% and 15%) which indicates that this gene cluster could be a house-keeping gene cluster for *Methanohalobium* to perform osmo-adaptation and regulation. The involvement of glycine betaine in strain ERBA was also confirmed by a growth experiment with 0.5mM glycine betaine supplement to standardized growth media at 15% salinity. Cells grown with glycine betaine showed significantly shorter lag phase (by 3 days), doubled growth yield and increased the amount of methane produced.

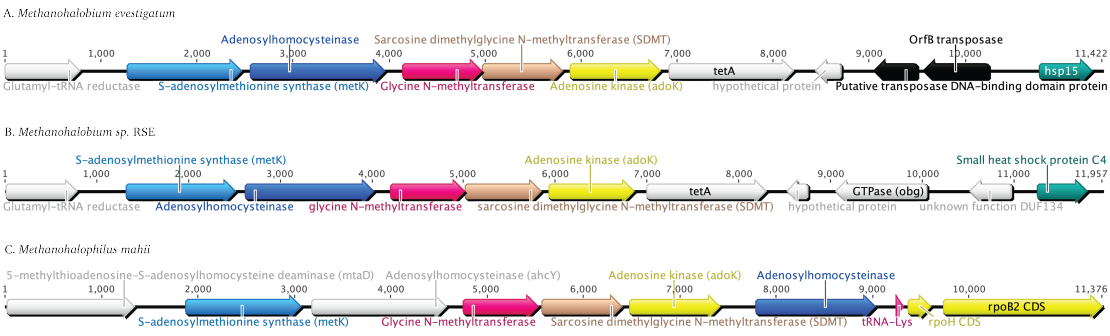


Figure 14. Conserved gene clusters for betaine *de novo* biosynthesis in *Methanohalobium* and *Methanohalophilus*. Genes for S-adenosylmethionine synthases are in light blue, adenosylhomocysteinases are in dark blue, glycine

sarcosine methyltransferases are in red, sarcosine dimethylglycine methyltransferases are in pink, and adenosine kinases are in yellow.

However, *Methanohalobium* do not have lysine-2,3-aminomutase (*ablA*) and β-lysine acetyltransferase (*ablB*), which are responsible for catalyzing *N*ε-acetyl-β-lysine formation from α-lysine, an osmolyte commonly found in methanogens (such as Methanohalophilus species, *Methanococcoides* species, *Methanosarcina mazei* Gö1, *Methanosarcina acetivorans*, *Methanosarcina barkeri*, *Methanococcus jannaschii*, and *Methanococcus maripaludis*). Candidate genes for the synthesis of ectoine, trehalose, and 2-sulfotrehalose were not detected either.

2.3.12 Antioxidant enzymes and oxygenases

Antioxidant enzymes hold the key for methanogens to thrive when exposed to environments with minimal concentrations of oxygen. Multiple enzymes related to scavenging oxidants are predicted from both *Methanohalobium* genomes. Notably, these genomes carry amongst the highest amounts of thioredoxin homologs of all the methanogens. Similar to *Methanocaldococcus jannaschii* (Susanti *et al.*, 2014), thioredoxin proteins participate in protecting multiple cellular processes against oxygen in *Methanohalobium*. Much to our surprise, *Methanohalobium* lack catalase and catalase-peroxidase, which are robust and commonly used by *Methanosarcinales* methanogens to detoxify H_2O_2 (Spring *et al.*, 2010).

It is very interesting that *Methanohalobium* possess predicted carotenoid oxygenases (ERBA_01937 and Metev_0061) belonging to the RPE65

superfamily retinal pigment epithelial membrane protein. They appear to be homologous to the bacterial carotenoid oxygenase (at the amino acid level, 45% identity to the homologue in *Geobacter sulfurreducens* and 30% identity to the closest characterized homolog, an apocarotenoid-15,15'-oxygenase in *Synechocystis* sp. PCC 6803 (Ruch *et al.*, 2005)). The cleavage products of this enzyme are retinal and corresponding retinal derivatives that serve in crucial biological functions, such as energy and signal transducers (Spudich *et al.*, 2000). The signaling function of this protein in *Methanohalobium* is definitely worth future experimental investigations.

2.3.13 Mobility and chemotaxis

Although motility is not observed in cells grown in laboratory conditions, both genomes showed a capacity for motility via archaellum (former: archaeal flagellum (Jarrell and Albers, 2012)). They both contain large numbers of genes (many of which appear in clusters) for archaellum biosynthesis and chemotaxis (20 genes in ERBA and 28 genes in *Mhb. evestigatum*). We speculate that *Methanohalobium* might mainly utilize archaella to swim around looking for nutrients and the expression and assembly of archaella in *Methanohalobium* could be induced by nutrient limiting conditions in their natural habitats. It is not clear if *Methanohalobium* could produce gas vesicles to give cells buoyancy. Except for one gene (ERBA_00062 annotated as a CbbQ/NirQ/NorQ domain protein) containing a gas vesicle GvpN domain, other proteins that are associated with the production of gas vesicles have not been identified.

2.3.14 Membrane transporters

Methanohalobium have the genomic capacity to synthesize several ATP-binding cassette transporters (ABC transporters) for translocating substrates across cell membranes by utilizing the energy from ATPs. Both genomes encode transporters for tungstate/molybdate, phosphate, sulfate, sulfonate/nitrate/taurine, glycine betaine/proline, iron, zinc, cobalt, nickel, multidrug, lipo-oligosaccharide, lipoprotein, and heme. Genes for peptide/nickel ABC transporter are only identified in the genome of ERBA (as already mentioned in the section on Sulfur metabolism). *Methanohalobium* could use their amino acid permease (PotE) and putative amino acid-proton symporter (YbeC) for exogenous amino acid acquisition. In addition to the diverse membrane transporters for inorganic compounds involved in osmoadaptation mentioned in the section of Osmoadaptation, genes that encode uptake systems for sodium/proline (PutP), magnesium/cobalt (CorA), zinc (ZupT), and divalent cation (MgtE) are also found in both genomes. They also have biotin transporters to confer the ability to take up biotin.

To help in protecting the cell from electrophiles, heavy metals and toxic compounds, *Methanohalobium* seem to utilize kebB type K^+ efflux glutathione-regulated potassium efflux systems, copper/silver exporters (Cop), macrolide exporters, cation efflux pump (Fief), arsenic transporters, resistance-nodulation-cell division (RND) transporters and multi-drug exporters to confer resistance.

2.3.15 Lipid biosynthesis

The membrane lipids of archaea consist of isoprenoid alkyl chains, which are attached to glycerol moieties via stable ether linkages (Villanueva *et al.*, 2014). The following components have been identified in *Mhb. evestigatum* (Koga and Nakano, 2008): caldarchaeol and cyclic archaeol as core lipid, D-glucose as glycolipid sugar, and *myo*-inositol, ethanolamine and glycerol as phospholipid-polar head groups. Due to the presence of cyclic archaeol and caldarchaeol, the combination of lipid component parts (LCP) is different from other members of the family *Methanosarcinaceae* (Koga and Nakano, 2008). The genomes of *Methanohalobium* harbor corresponding enzymes proposed for the lipid biosynthesis pathway in *Methanococcoides burtonii* (Allen *et al.*, 2009),

2.3.16 Prophage and CRISPR/Cas systems

The two species have distinct signatures of viral defense and possible prophage integration. The CRISPR (clustered regularly interspaced short palindromic repeats)/Cas system preserves invading genetic material and incorporates it into surveillance complexes to achieve adaptive immunity (Makarova *et al.*, 2011). The system consists of spacers (DNA fragments inserted from invading foreign DNAs), repeats and CRISPR-associated proteins. No CRISPR/Cas array has been identified in the complete genome of *Mhb. evestigatum*. However, 3 incomplete putative prophage regions ranging from 9 Kb to 11.2Kb have been identified in this genome according to PHAST. In contrast, no prophage regions have been identified in the ERBA draft genome. The *Mhb.* strain ERBA draft genome comprises two CRISPR loci and a total of 86 spacers. CRISPR locus 1 (3594 bp) is flanked by Cas protein coding gene arrays *cas3-cas5-cas7-*

cas8a1 belonging to Cas type A-I. CRISPR locus 2 (2237 bp) is not associated with any Cas protein coding genes. The two CRISPR loci use different repeats classified via CRISPRmap as motif 13-type and superclass F/class7-type repeat. Two repetitive spacers and four self-targeting spacers (targeting genes for molecular chaperone, DUF2124 hypothetical protein, inosine-5'-monophosphate dehydrogenase and ATPase) are observed. Thus, we believe that the CRISPR/Cas system does not only hold the potential for ERBA to survive when exposed to foreign genetic elements in the deep-sea hypersaline environment of the Red Sea, but may also have a role to play in gene regulation or autoimmunity.

2.3.17 Linking strain ERBA to the deep-sea hypersaline brine environment

We have identified a number of reasons that could explain why strain ERBA could thrive in the polyextreme habitat of Erba Deep, the Red Sea. Firstly, it performs methylotrophic methanogenesis that is thermodynamically much more favorable than aceticlastic and hydrogenotrophic methanogenesis (Oren, 2011). In order to reinforce this process under deep-sea hypersaline brines, genes for essential enzymes involved in methanogenesis and methanogenesis-associated electron transfer are present in multiple copies, many of which are species-specific genes in strain ERBA. Secondly, the genome of ERBA indicates that it has the potential to utilize more diverse approaches for nutrient acquisition and assimilation. This has been demonstrated by the diverse and unique membrane transporters and metabolic pathways analyzed throughout the chapter. Thirdly, the ERBA genome reveals diverse strategies for osmo-response through its membrane channels, ABC transporters, and *de novo*

biosynthesis. Fourthly, to better transduce external signals into cells in order to generate physiological response in the deep-sea, strain ERBA carries 20 unique genes that could be involved in the signal transduction apparatuses. Fifthly, the niche adaptation of ERBA could also be reflected by its unique cell wall/membrane/envelope signatures. Several unique genes for glycosyltransferases are present in ERBA indicating potential differences in surface layer glycoproteins in these two species. Three unique copies of the *tolB* gene suggest that ERBA could employ the Tol-Pal protein cell envelope system to maintain maintenance of cell envelope integrity during cell division. The corresponding enzymes in *E. coli* have shown to be required for cell growth at high hydrostatic pressure (Black *et al.*, 2013). Finally, the genome of ERBA indicates that it is well prepared for conducting basic cellular processes under high salinity. Unique components in chromosome condensation and folding, DNA replication, recombination and repair, transcription regulation, tRNA-queuosine biosynthesis, posttranslational modification and chaperones are also found in the genome of ERBA. The unique genes in strain ERBA are also summarized in Table S1.

2.3.18 Salt-dependent expression profiling in strain ERBA

As already mentioned in the section on Cell morphology and characteristics, strain ERBA can grow in a board range of salinities (4% – 29%). In Erba Deep, where this strain was isolated, *Methanohalobium* populations have been detected both in the brine-seawater interface (at 8% – 10.6% salinity) and the brine underneath (at 18% salinity) which indicates that *Methanohalobium* would have to deal with physicochemical changes (especially salinity) in their

natural habitat. In order to gain insights into salinity-dependent transcriptional characteristics of strain ERBA, RNA-seq data was compared using strain ERBA cells grown at a variety of salinities (5% NaCl (low salinity), 10% NaCl (its optimal growth salinity) and 15% NaCl (high salinity, resembling the salinity observed in Erba Brine)). 0.5M – 0.7M paired-end reads after QC were used for Bowtie2 mapping and subsequent analysis (see also Table S2).

A multidimensional scaling plot (MDS plot) was used to examine the distances between NaCl-dependent gene expression profiles reflected by the $\log_2$-fold changes between the samples. In the plot shown in Figure 15, dimension 1 clearly separates the control (10% NaCl) samples from samples from ERBA cells that are grown under less favorable salinities (5% and 15%), while dimension 2 distinguishes the high salt samples (15% NaCl) from others. The 5% NaCl samples appear more heterogeneous than other samples. We observed good consistency in biological replicates and clear separation of gene expression change among the three salinities tested.

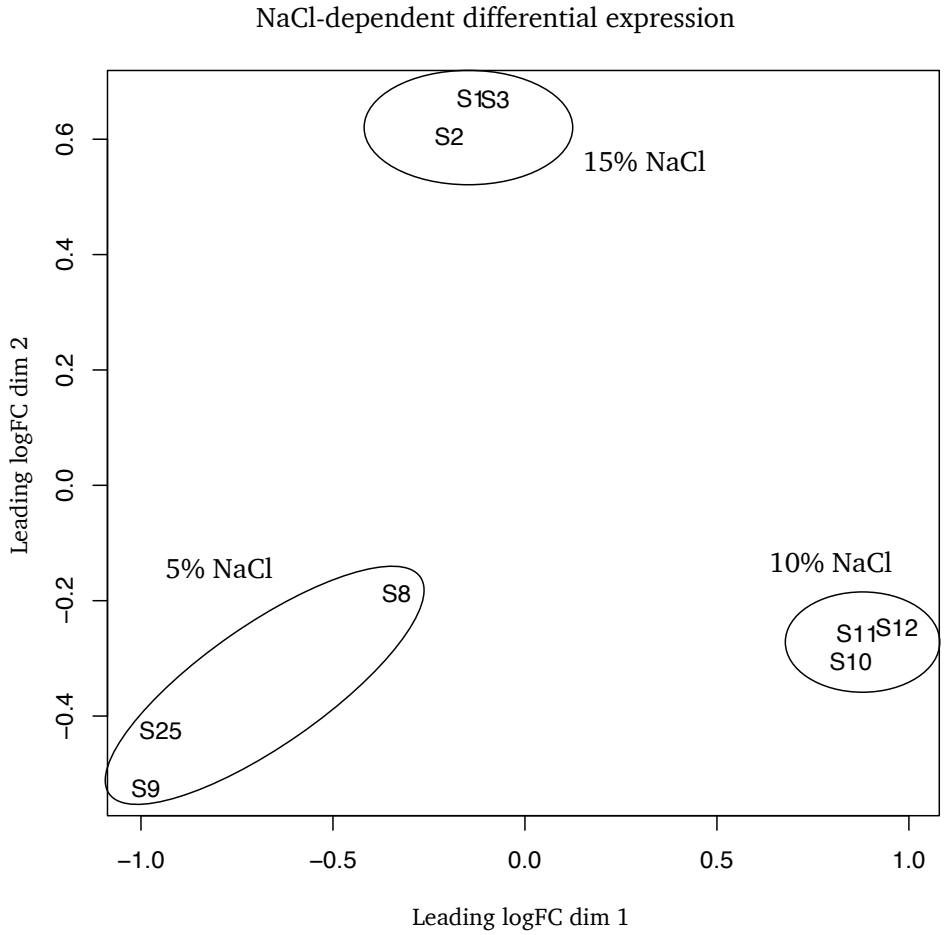


Figure 15. MDS plot of salt-dependent gene $\log_2$ fold changes in strain ERBA.

The global transcriptome analysis revealed that collectively more than 99% of the predicted genes of ERBA were expressed under at least one of the NaCl growth conditions (5%, 10% and 15%) examined. A large number of protein-encoding genes were highly expressed under different salinities (439 to 601 genes). 362 highly expressed “housekeeping gene” repertoires in ERBA were generated based on protein-encoding genes with FPKM values greater than 350 under all salinities tested (Table 12 and also listed in Table S3). Those genes which were highly expressed were mostly characterized as belonging to the groups Translation, ribosomal structure and biogenesis (78 genes), Energy

production and conversion (41 genes), and Coenzyme transport and metabolism (35 genes) (see also Figure 16), which indicates their fundamental role in maintaining cellular function in strain ERBA. In addition, all the genes coding for the bacterial type of the betaine *de novo* biosynthesis gene cluster previously mentioned in the section of Osmoadaptaion were all high expressed, suggesting the essential role of betaine in the halophilic lifestyle of ERBA.

Table 12. Summary of the RNA-seq data associated with genes in the *Methanohalobium* strain ERBA. The significance of differentially expressed genes is based on those with *p values* under 0.01 in the Fisher's Exact Test in edgeR.

	Number of genes	% with respect to the total number of genes of strain ERBA
M. ERBA total number of genes	2158[a]	-
Genes represented in RNA-seq	2143	99.3
Highly expressed protein-encoding genes		
Genes with FPKM value over 350 at 5% salinity	601	28.5
Genes with FPKM value over 350 at 10% salinity	439	20.8
Genes with FPKM value over 350 at 15% salinity	546	25.9
Genes with FPKM value over 350 at 5%, 10% and 15% salinities	362	17.2
Differentially expressed genes at 15% salinity relative to 10% salinity	563	26.1
Up-regulated genes at 15% salinity relative to 10% salinity	256	11.9
Down-regulated genes at 15% salinity relative to 10% salinity	307	14.2
Differentially expressed genes at 5% salinity relative to 10% salinity	793	36.7
Up-regulated genes at 5% salinity relative to 10% salinity	433	20.1
Down-regulated genes at 5% salinity relative to 10% salinity	360	16.7

[a] TAG codon interrupted mono-, di-, and trimethylamine methyltransferase genes were manually cured.

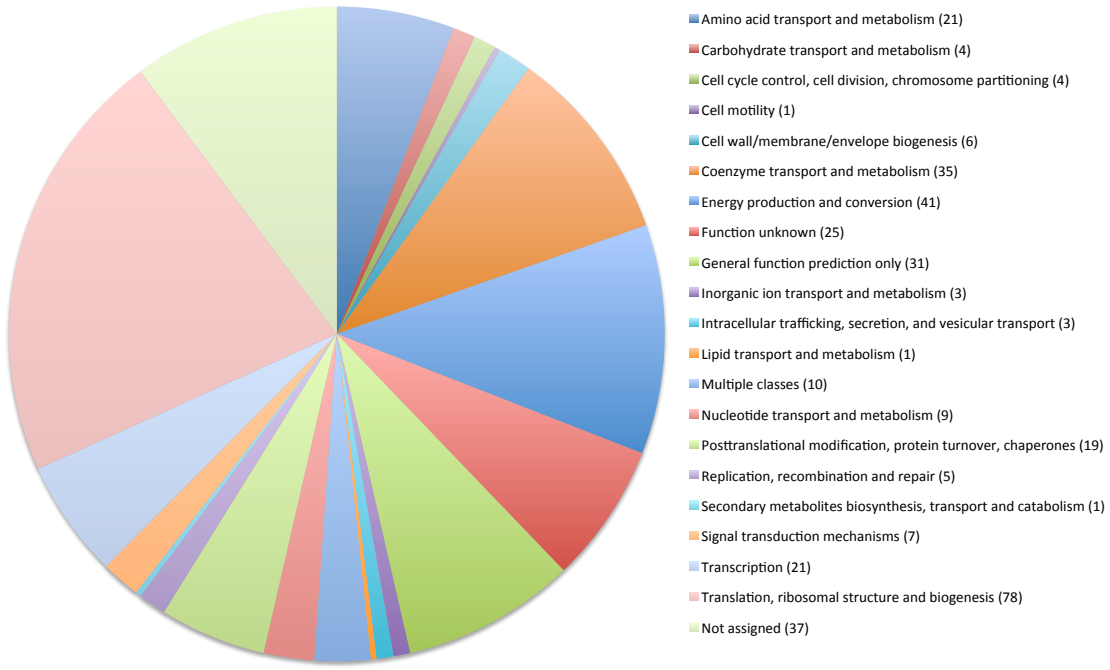


Figure 16. Distribution of housekeeping genes in COGs.

The gene expression in strain ERBA was significantly stimulated by changes in osmolarity away from its optimal salinity (10% NaCl). The abundance of transcripts for 563 genes (26.1% of the total predicted genes) was significantly (p *value* <0.01) differentially expressed at 15% salinity compared with 10%. Even higher numbers of significantly (p *value* <0.01) differentially expressed genes (793 genes, accounted for 36.7% of total predicted genes) was observed at 5% NaCl related to 10% (see also Table 12 and values in Table S4-S7). To identify functional gene categories characteristic of salt adaptation, these significantly regulated genes were classified against the COG database and the results are depicted in Figure 17. We observed that the shifts in salinity have resulted in significant changes in gene expression involving in multiple cell functions. Substantial amounts of these genes could not be assigned to any of the COG categories, or they could only be classified as belonging to General

function prediction only or to Function unknown, indicating that proteins with currently undefined functions are sensitive to the changes in salinity. Subsequently, COG categories of Energy production and conversion, Translation, ribosomal structure and biogenesis, Posttranslational modification, protein turnover, chaperones, and Coenzyme transport and metabolism comprise abundant numbers of significantly differentially expressed genes, revealing that these process could facilitate salt adaptation.

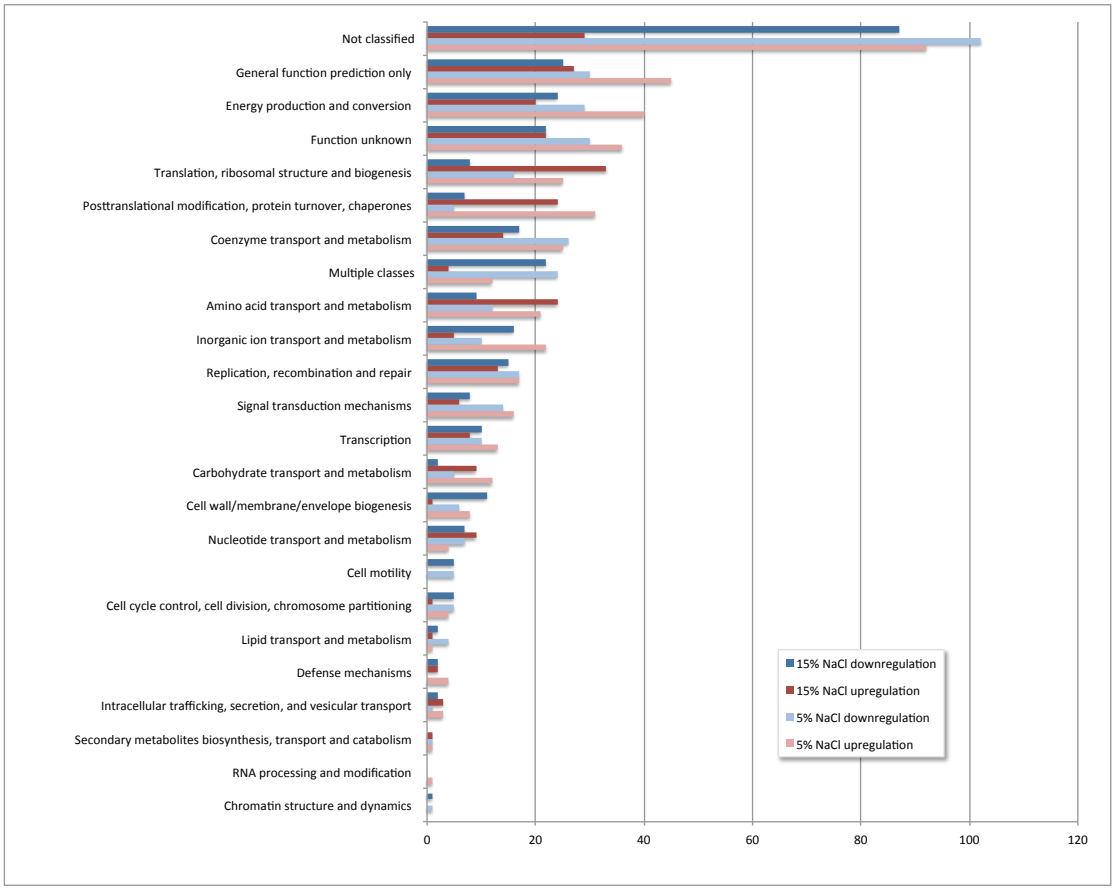


Figure 17. COG functional classification of significantly up or down regulated genes under different salinities in strain ERBA.

The high salt adaptation (15% NaCl) is mainly characterized by upregulation of the following genes. The genes for ferrous, zinc ion, and cobalamin

transport (ERBA_02147, ERBA_02145, and ERBA_00581) are all upregulated, which might be a result of higher demand as cofactors under high salt. For example, iron is an important component in iron-sulfur proteins involved in oxidation-reduction processes. The drastic increase in iron uptake in ERBA under 15% NaCl might be correlated with increased iron-binding proteins. Similarly, genes for glycine, lysine, glutamate, and arginine synthesis related hydroxylamine reductase (ERBA_00432), NADPH-dependent glutamate synthase (ERBA_01098 and ERBA_01099) and N-acetylglutamate synthase (ERBA_01116) are upregulated as well. Glycine and lysine are precursors of compatible solutes glycine betaine and Nε-acetyl-β-lysine which are commonly found in methanogens, while glutamate and arginine themselves are osmoprotectants (Xu *et al.*, 2011; M C Lai *et al.*, 1991). In combination with the observation that genes for betaine *de novo* synthesis (ERBA_00872 – ERBA_00868) are constantly highly expressed in ERBA, it indicates the important role that compatible solutes play in helping ERBA adapt to high salt. Genes related to translation, ribosomal structure, and biogenesis were also more abundant. Large members of ribosomal protein subunits and enzymes mediating amino acid incorporation into proteins were upregulated at 15% NaCl. 24 genes involved in posttranslational modification, protein turnover and chaperones processes were upregulated at 15% NaCl, including chaperone protein ClpB 2 (ERBA_00404), chaperone protein DnaK and DnaJ (ERBA_00091 and ERBA_00092), protein GrpE (ERBA_00090), proteasome (ERBA_00248) and thermosome proteins (ERBA_01590 and ERBA_01706).

ERBA seems to upregulate both protein folding and degradation processes to ensure protein quality at high salt.

In contrast, at 15% salinity, ERBA downregulated genes for methylotrophic methanogenesis, membrane bound electron transfer Rnf complex, V-type ATP synthase and NADH-quinone oxidoreductase, nitrogenases, chemotaxis and archaella, and specific chromosome partitioning and DNA-binding proteins. It is very interesting to see that the MinD chromosome partitioning ATPase gene (ERBA_02218), Helix-turn-helix family DNA binding protein encoding gene (ERBA_02219) and Non-histone chromosomal protein MC1 (ERBA_01556) were all strongly downregulated at 15% NaCl (greater than 4-fold). Their protein products function as cell division/septum formation inhibitor-activator (de Boer *et al.*, 1991), transcriptional repressors (Aggarwal *et al.*, 1988) and DNA bending and supercoiling (Laine *et al.*, 1991), which suggests that ERBA facilitate transcription and replication by reducing the expression of these inhibitory genes under elevated salinities.

A similar general trend in gene expression was also observed under 5% salinity for genes involved in inorganic ion transport (enhanced expression), methylotrophic methanogenesis, membrane bound electron transfer Rnf complex, V-type ATP synthase and NADH-quinone oxidoreductase, nitrogenases, and archaella (reduced expression). In addition, to adapt to this relatively low osmotic condition (5% NaCl), ERBA exhibited enhanced gene expression for Trk system potassium uptake proteins (TrkAH), sodium antiporter and MscS mechanosensitive ion channel, while at the same time

significantly decreased the expression of genes for glycine betaine uptake (ERBA_00742, ERBA_00743 and ERBA_00906).

Overall, our data suggest that methanogenesis and its related energy conserving electron-transfer in ERBA are only highly expressed around its optimal growth salinity (10% NaCl), which corresponds to both highest growth and methane production, indicating that these cell machineries are well adapted to this moderate salinity. For this halophilic methanogen, lowering its environmental salinity does not necessarily favor increased energy metabolism. To depict the abundances of genes involved in this important energy-producing process, FPKM values of all genes predicted to participate in methanogenesis in ERBA were clustered across samples and plotted in the heatmap shown in Figure 18. Consistent with a previous study on cold-adapted methanogens (Campanaro *et al.*, 2011), the main genes involved in methanogenesis are generally upregulated at optimal growth conditions as shown in the upper cluster of the horizontal dendogram (c.f. Figure 18). As previously noted, methylotrophic methanogens usually carry multiple copies of genes involved in substrate-specific methyltransfer (Galagan *et al.*, 2002; Allen *et al.*, 2009). Despite being at low transcription abundances, multiple species-specific genes in ERBA exhibited elevated expression when not at the optimal salinity (10% NaCl), indicating their potential role in compensating for the downregulation of main genes involved in the same/or similar functions of methanogenesis in order to cope with osmotic stress.

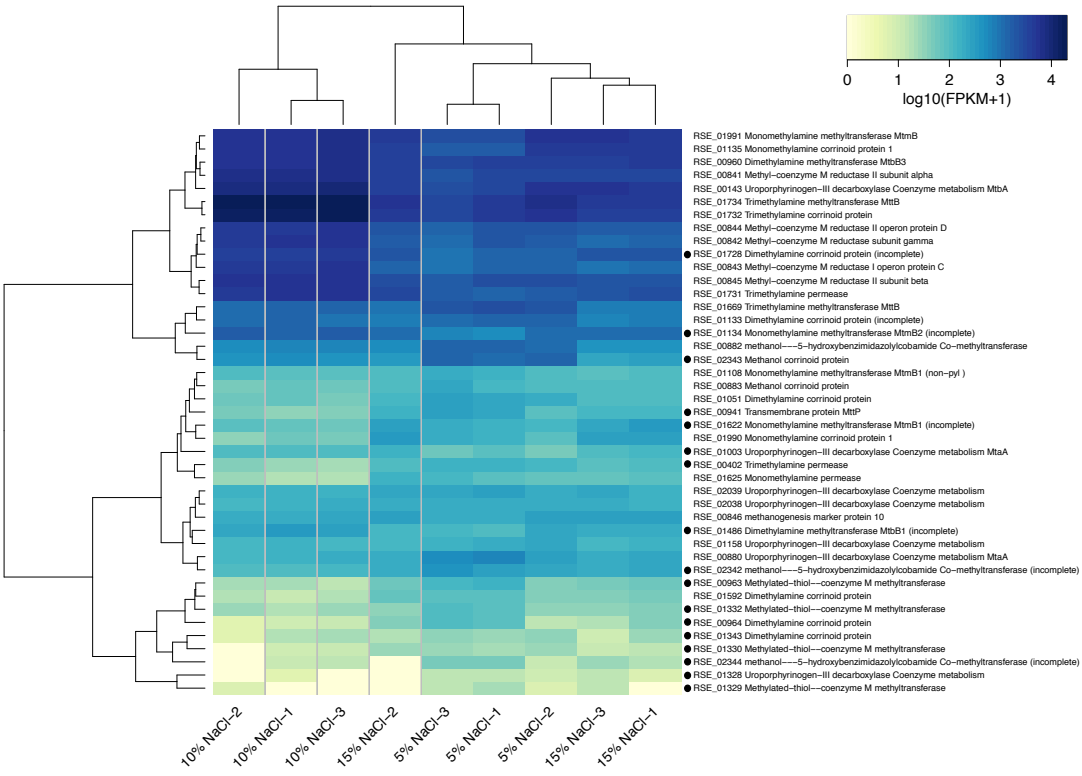


Figure 18. Gene expression profiles of genes involved in methanogenesis in samples under 5%, 10% and 15% NaCl. The horizontal dendogram shows genes that are annotated on the right side of corresponding row. The vertical dendogram shows RNA samples. Both dendograms were generated by hierarchical clustering. Circles next to gene IDs indicate species-specific genes in ERBA.

2.4 Conclusions

Based on physiological, genomic and transcriptomic characterizations and comparisons, we illustrated distinct lifestyles within the halophilic methanogen genus *Methanohalobium*. Their genomic traits have demonstrated their remarkable adaptation and metabolic capacities. Multiple cellular processes and a large number of genes in the deep-sea methanogen are involved in coping with unfavorable salinities at the transcriptome level. The current work not only illustrates some of the important characteristics of this poorly understood genus of methanogenic archaea, but it also helps to bridge the gap in our understanding in halophiles.

2.5 Supplementary materials

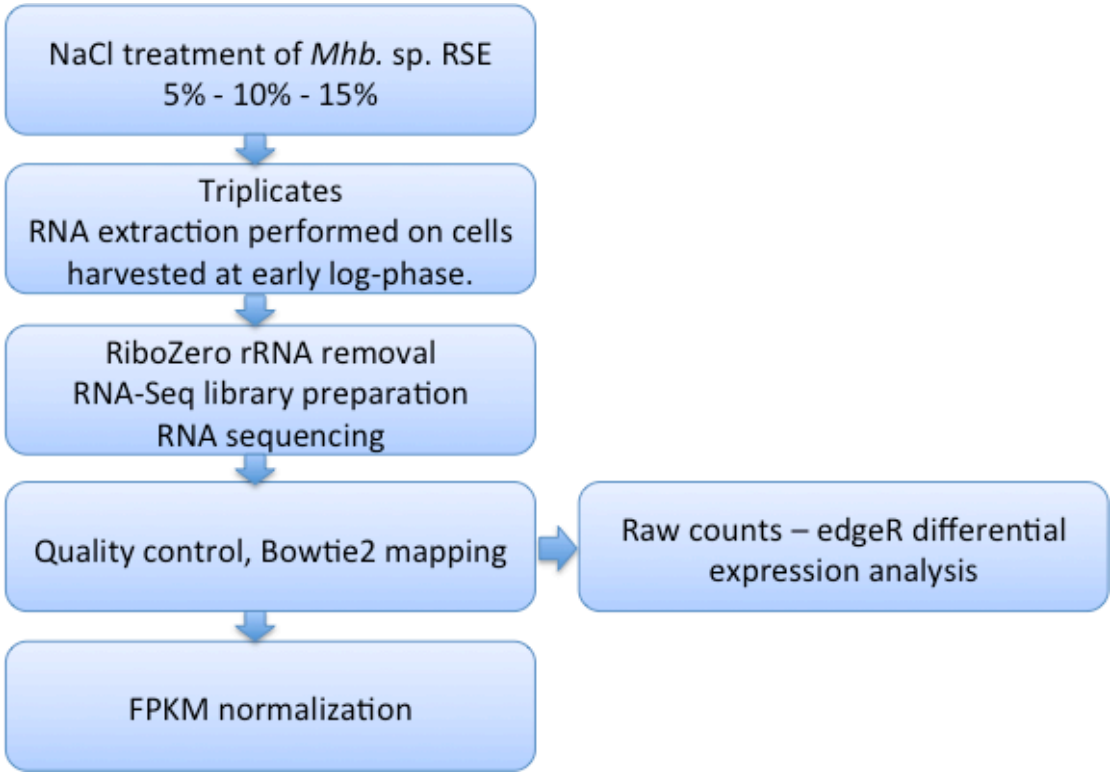


Figure S1. Overall schematic diagram of methods used in the RNA-seq of strain ERBA.

Table S1. Unique genes in strain ERBA.

Gene ID	INDIGO annotation	COG classification	KEGG classification
ERBA_00069	Putative threonine-phosphate decarboxylase	Amino acid transport and metabolism	K04720:threonine-phosphate decarboxylase [EC:4.1.1.81]
ERBA_00611	Oligopeptide transport ATP-binding protein OppF	Amino acid transport and metabolism	K10823:oligopeptide transport system ATP-binding protein
ERBA_00612	Dipeptide transport ATP-binding protein DppD	Amino acid transport and metabolism	K15583:oligopeptide transport system ATP-binding protein
ERBA_00613	Dipeptide transport system permease protein DppC	Amino acid transport and metabolism	K15586:nickel transport system permease protein
ERBA_00614	Glutathione transport system permease protein GsiC	Amino acid transport and metabolism	K13890:glutathione transport system permease protein
ERBA_00616	Glutathione-binding protein GsiB	Amino acid transport and metabolism	K03099:son of sevenless
ERBA_00806	putative amino acid-proton symporter YbeC	Amino acid transport and metabolism	K13863:solute carrier family 7 (cationic amino acid transporter), member 1
ERBA_00982	Putative threonine-phosphate decarboxylase	Amino acid transport and metabolism	K04720:threonine-phosphate decarboxylase [EC:4.1.1.81]
ERBA_01066	putative aspartoacylase	Amino acid transport and metabolism	K15784:N-alpha-acetyl-L-2,4-diaminobutyrate deacetylase [EC:3.5.1.-]
ERBA_01389	nemo like kinase	Amino acid transport and metabolism	K04468:nemo like kinase [EC:2.7.11.24]
ERBA_01548	putative cysteine desulfurase	Amino acid transport and metabolism	K04487:cysteine desulfurase [EC:2.8.1.7]
ERBA_02293	GCN5-related N-acetyltransferase	Amino acid transport and metabolism	K02961:small subunit ribosomal protein S17
ERBA_00272	23-bisphosphoglycerate-independent phosphoglycerate mutase	Carbohydrate transport and metabolism	K01834:2,3-bisphosphoglycerate-dependent phosphoglycerate mutase [EC:5.4.2.11]
ERBA_00402	trimethylamine permease	Carbohydrate transport and metabolism	
ERBA_00941	transmembrane protein MttP	Carbohydrate transport and metabolism	
ERBA_01275	Pyruvate kinase	Carbohydrate transport and metabolism	K00873:pyruvate kinase [EC:2.7.1.40]
ERBA_01336	Aquaporin AqpM	Carbohydrate transport and metabolism	
ERBA_01490	Citrate-proton symporter	Carbohydrate transport and metabolism	K02871:large subunit ribosomal protein L13
ERBA_01589	4Fe-4S ferredoxin iron-sulfur binding domain containing protein	Carbohydrate transport and metabolism	
ERBA_02345	Pyruvate kinase	Carbohydrate transport and metabolism	K00873:pyruvate kinase [EC:2.7.1.40]
ERBA_00294	Protein CrcB	Cell cycle control, cell division, chromosome partitioning	
ERBA_00295	Protein CrcB	Cell cycle control, cell division, chromosome partitioning	
ERBA_01515	type II secretion system protein E	Cell motility	K03196:type IV secretion system protein VirB11

ERBA_01516	type II secretion system protein	Cell motility	
ERBA_01248	glycogen synthase	Cell wall/membrane/envelope biogenesis	K16153:phosphorylase / glycogen(starch) synthase [EC:2.4.1.1 2.4.1.11]
ERBA_01278	glycoprotein 3-alpha-L-fucosyltransferase	Cell wall/membrane/envelope biogenesis	K00754:
ERBA_01280	Undecaprenyl-phosphate 4-deoxy-4-formamido-L-arabinose transferase	Cell wall/membrane/envelope biogenesis	K00754:
ERBA_01281	Glycosyltransferase Cell envelope biogenesi	Cell wall/membrane/envelope biogenesis	K00754:
ERBA_01288	Bifunctional protein GlmU	Cell wall/membrane/envelope biogenesis	K04042:bifunctional UDP-N-acetylglucosamine pyrophosphorylase / Glucosamine-1-phosphate N-acetyltransferase [EC:2.7.7.23 2.3.1.157]
ERBA_01792	MscS mechanosensitive ion channel	Cell wall/membrane/envelope biogenesis	K00606:3-methyl-2-oxobutanoate hydroxymethyltransferase [EC:2.1.2.11]
ERBA_02014	Small-conductance mechanosensitive channel	Cell wall/membrane/envelope biogenesis	K01834:2,3-bisphosphoglycerate-dependent phosphoglycerate mutase [EC:5.4.2.11]
ERBA_00074	Biotin synthase	Coenzyme transport and metabolism	K01012:biotin synthase [EC:2.8.1.6]
ERBA_00281	von Willebrand factor type A	Coenzyme transport and metabolism	K05027:calcium-activated chloride channel regulator 1
ERBA_00289	putative tRNA sulfurtransferase	Coenzyme transport and metabolism	K03151:thiamine biosynthesis protein ThiI
ERBA_00615	Ubiquinone-menaquinone biosynthesis methyltransferase ubiE	Coenzyme transport and metabolism	K00599:
ERBA_00640	Methyl-coenzyme M reductase II subunit beta	Coenzyme transport and metabolism	K11529:glycerate 2-kinase [EC:2.7.1.165]
ERBA_00963	Methylated-thiol--coenzyme M methyltransferase	Coenzyme transport and metabolism	K01599:uroporphyrinogen decarboxylase [EC:4.1.1.37]
ERBA_01003	Uroporphyrinogen-III decarboxylase Coenzyme metabolism	Coenzyme transport and metabolism	K14080:[methyl-Co(III) methanol-specific corrinoid protein]:coenzyme M methyltransferase [EC:2.1.1.246]
ERBA_01067	Ribosomal protein S6 modification protein	Coenzyme transport and metabolism	K14940:gamma-F420-2:alpha-L-glutamate ligase [EC:6.3.2.32]
ERBA_01328	Uroporphyrinogen-III decarboxylase Coenzyme metabolism	Coenzyme transport and metabolism	K14080:[methyl-Co(III) methanol-specific corrinoid protein]:coenzyme M methyltransferase [EC:2.1.1.246]
ERBA_01329	Methylated-thiol--coenzyme M methyltransferase	Coenzyme transport and metabolism	K14080:[methyl-Co(III) methanol-specific corrinoid protein]:coenzyme M methyltransferase [EC:2.1.1.246]
ERBA_01330	Methylated-thiol--coenzyme M methyltransferase	Coenzyme transport and metabolism	K14080:[methyl-Co(III) methanol-specific corrinoid protein]:coenzyme M methyltransferase [EC:2.1.1.246]
ERBA_01332	Methylated-thiol--coenzyme M methyltransferase	Coenzyme transport and metabolism	K01599:uroporphyrinogen decarboxylase [EC:4.1.1.37]
ERBA_01334	putative cyclic pyranopterin monophosphate synthase	Coenzyme transport and metabolism	K03639:molybdenum cofactor biosynthesis protein
ERBA_01335	phenylacetate--CoA ligase	Coenzyme transport and metabolism	K01912:phenylacetate-CoA ligase [EC:6.2.1.30]
ERBA_01429	putative cobalt-precorrin-3B C-methyltransferase	Coenzyme transport and metabolism	K04720:threonine-phosphate decarboxylase [EC:4.1.1.81]
ERBA_01669	Trimethylamine methyltransferase MttB2	Coenzyme transport and metabolism	K14083:trimethylamine---corrinoid protein Co-methyltransferase [EC:2.1.1.250]
ERBA_01733	Trimethylamine methyltransferase MttB	Coenzyme transport and metabolism	K14083:trimethylamine---corrinoid protein Co-methyltransferase [EC:2.1.1.250]
ERBA_01837	putative molybdenum-pterin-binding protein	Coenzyme transport and metabolism	K02017:molybdate transport system ATP-binding protein [EC:3.6.3.29]
ERBA_02341	putative cobalt-precorrin-3B C-methyltransferase	Coenzyme transport and metabolism	K04720:threonine-phosphate decarboxylase [EC:4.1.1.81]
ERBA_00720	CRISPR-associated protein CXXC CXXC region	CRISPR-Cas system-associated protein Cas8a1 CRISPR and associated Cas proteins comprise a system for heritable host defense by prokaryotic cells against phage and other foreign DNA: Large proteins	
ERBA_01385	ISA1083-2 transposase	DDE superfamily endonuclease	
ERBA_01486	Dimethylamine methyltransferase MtbB1	Dimethylamine methyltransferase It demethylates dimethylamine	K16178:dimethylamine---corrinoid protein Co-methyltransferase [EC:2.1.1.249]
ERBA_00961	Dimethylamine methyltransferase MtbB2	dimethylamine:corrinoid methyltransferase	K16178:dimethylamine---corrinoid protein Co-methyltransferase [EC:2.1.1.249]
ERBA_00639	family 2 glycosyl transferase	DPM_DPG-synthase_like is a member of the Glycosyltransferase 2 superfamily DPM1 is the catalytic subunit of eukaryotic dolichol-phosphate mannose is a transmembrane-bound enzyme of the endoplasmic reticulum involved in protein N-linked glycosylation	K00721:dolichol-phosphate mannosyltransferase [EC:2.4.1.83]
ERBA_00310	CoB--CoM heterodisulfide reductase iron-sulfur subunit A 1	Energy production and conversion	K03522:electron transfer flavoprotein alpha subunit
ERBA_00744	4Fe-4S ferredoxin iron-sulfur binding domain containing protein	Energy production and conversion	K00390:phosphoadenosine phosphosulfate reductase [EC:1.8.4.8]
ERBA_01172	Plastocyanin	Energy production and conversion	K00368:nitrite reductase (NO-forming) [EC:1.7.2.1]
ERBA_01247	Galactose-1-phosphate uridylyltransferase	Energy production and conversion	K00965:UDPglucose--hexose-1-phosphate uridylyltransferase [EC:2.7.7.12]
ERBA_01331	NADH-dependent flavin oxidoreductase	Energy production and conversion	K10797:2-enoate reductase [EC:1.3.1.31]
ERBA_01341	putative aldehyde dehydrogenase AldX	Energy production and conversion	K00128:aldehyde dehydrogenase (NAD+) [EC:1.2.1.3]
ERBA_01347	CoB--CoM heterodisulfide reductase iron-sulfur subunit A	Energy production and conversion	K00338:NADH-quinone oxidoreductase subunit I [EC:1.6.5.3]
ERBA_01549	Nitrogen fixation protein NifU	Energy production and conversion	K00566:tRNA-specific 2-thiouridylase [EC:2.8.1.-]
ERBA_02067	Nitroreductase Energy production	Energy production and conversion	K01212:levanase [EC:3.2.1.65]
ERBA_01507	RDD domain-containing protein	Function unknown	
ERBA_01547	NAD-dependent dehydrogenase	Function unknown	K04085:tRNA 2-thiouridine synthesizing protein A [EC:2.8.1.-]
ERBA_01718	Secreted effector protein pipB2	Function unknown	K08873:PI-3-kinase-related kinase SMG-1
ERBA_00308	transcriptional regulator protein	General function prediction only	

ERBA_00393	NADPH-dependent FMN reductase protein	General function prediction only	K00299:FMN reductase [EC:1.5.1.38]
ERBA_00430	MMPL domain containing protein	General function prediction only	K07787:Cu(I)/Ag(I) efflux system membrane protein CusA/SilA
ERBA_00535	beta-lactamase	General function prediction only	
ERBA_00717	Putative CRISPR-associated nuclease-helicase Cas3	General function prediction only	K00595:precorrin-6Y C5,15-methyltransferase (decarboxylating) [EC:2.1.1.132]
ERBA_00807	2-succinyl-6-hydroxy-2 4-cyclohexadiene-1-carboxylate synthase	General function prediction only	K08680:2-succinyl-6-hydroxy-2,4-cyclohexadiene-1-carboxylate synthase [EC:4.2.99.20]
ERBA_00904	NADPH-dependent FMN reductase protein	General function prediction only	K00299:FMN reductase [EC:1.5.1.38]
ERBA_00939	Phosphoglycolate phosphatase	General function prediction only	K01558:wingless-type MMTV integration site family, member 16
ERBA_00964	Dimethylamine corrinoid protein 2	General function prediction only	K16179:dimethylamine corrinoid protein
ERBA_01159	polyketide cyclase-like protein	General function prediction only	
ERBA_01162	acetyltransferase family protein	General function prediction only	K00680:
ERBA_01277	putative macrolide acetyltransferase	General function prediction only	K09691:lipopolysaccharide transport system ATP-binding protein
ERBA_01322	PKD domain containing protein	General function prediction only	K01384:wingless-type MMTV integration site family, member 11
ERBA_01343	Dimethylamine corrinoid protein 2	General function prediction only	K16179:dimethylamine corrinoid protein
ERBA_01397	hydrolase	General function prediction only	K01055:3-oxoadipate enol-lactonase [EC:3.1.1.24]
ERBA_01479	oxidoreductase molybdopterin binding domain containing protein	General function prediction only	K00387:sulfite oxidase [EC:1.8.3.1]
ERBA_01540	Phosphodiesterase YfcE	General function prediction only	K01525:bis(5'-nucleosyl)-tetraphosphatase (symmetrical) [EC:3.6.1.41]
ERBA_01728	Dimethylamine corrinoid protein 2	General function prediction only	K16179:dimethylamine corrinoid protein
ERBA_01756	NADPH-dependent FMN reductase protein	General function prediction only	
ERBA_02068	Putative mycofactocin radical SAM maturase MftC	General function prediction only	K03639:molybdenum cofactor biosynthesis protein
ERBA_02306	metallophosphoesterase	General function prediction only	
ERBA_02343	Methylated-thiol--corrinoid protein MtsB	General function prediction only	K14081:methanol corrinoid protein
ERBA_00020	Phosphate import ATP-binding protein PstB	Inorganic ion transport and metabolism	K02036:phosphate transport system ATP-binding protein [EC:3.6.3.27]
ERBA_00064	Arsenite resistance protein ArsB	Inorganic ion transport and metabolism	
ERBA_00095	Trk system potassium uptake protein TrkH	Inorganic ion transport and metabolism	K11643:chromodomain-helicase-DNA-binding protein 4 [EC:3.6.4.12]
ERBA_00096	Trk system potassium uptake protein TrkH	Inorganic ion transport and metabolism	K11643:chromodomain-helicase-DNA-binding protein 4 [EC:3.6.4.12]
ERBA_00343	Ktr system potassium uptake protein C	Inorganic ion transport and metabolism	K03426:NAD+ diphosphatase [EC:3.6.1.22]
ERBA_00344	Ktr system potassium uptake protein C	Inorganic ion transport and metabolism	K03426:NAD+ diphosphatase [EC:3.6.1.22]
ERBA_00384	Trk system potassium uptake protein TrkH	Inorganic ion transport and metabolism	K11643:chromodomain-helicase-DNA-binding protein 4 [EC:3.6.4.12]
ERBA_01073	sodium-calcium exchanger membrane protein	Inorganic ion transport and metabolism	K13749:solute carrier family 24 (sodium/potassium/calcium exchanger), member 1
ERBA_01147	sodium-calcium exchanger protein	Inorganic ion transport and metabolism	K16786:energy-coupling factor transport system ATP-binding protein [EC:3.6.3.-]
ERBA_01245	periplasmic copper-binding protein	Inorganic ion transport and metabolism	K01051:pectinesterase [EC:3.1.1.11]
ERBA_01297	Cadmium cobalt	Inorganic ion transport and metabolism	K14688:solute carrier family 30 (zinc transporter), member 1
ERBA_01344	Arsenite resistance protein ArsB	Inorganic ion transport and metabolism	
ERBA_01390	Trk system potassium uptake protein TrkH	Inorganic ion transport and metabolism	K11643:chromodomain-helicase-DNA-binding protein 4 [EC:3.6.4.12]
ERBA_01391	Trk system potassium uptake protein TrkA	Inorganic ion transport and metabolism	
ERBA_01510	Cobalamin import ATP-binding protein BtuD	Inorganic ion transport and metabolism	K02013:iron complex transport system ATP-binding protein [EC:3.6.3.34]
ERBA_01511	Cobalamin import system permease protein BtuC	Inorganic ion transport and metabolism	K02015:iron complex transport system permease protein
ERBA_01512	ligand binding protein	Inorganic ion transport and metabolism	K02016:iron complex transport system substrate-binding protein
ERBA_01562	putative calcium-transporting ATPase	Inorganic ion transport and metabolism	K01539:sodium/potassium-transporting ATPase subunit alpha [EC:3.6.3.9]
ERBA_01993	K+-dependent Na+-Ca+ exchanger related-protein	Inorganic ion transport and metabolism	K16786:energy-coupling factor transport system ATP-binding protein [EC:3.6.3.-]
ERBA_02214	nickel transport complex NikM subunit transmembrane	Inorganic ion transport and metabolism	K02009:cobalt/nickel transport protein
ERBA_01129	Protein TolB	Intracellular trafficking, secretion, and vesicular transport	K05685:macrolide transport system ATP-binding/permease protein [EC:3.6.3.-]
ERBA_02259	Protein TolB	Intracellular trafficking, secretion, and vesicular transport	K05685:macrolide transport system ATP-binding/permease protein [EC:3.6.3.-]
ERBA_02260	Protein TolB	Intracellular trafficking, secretion, and vesicular transport	K01423:
ERBA_01340	3-oxoacyl-acyl-carrier-protein reductase FabG	Lipid transport and metabolism	K00059:3-oxoacyl-[acyl-carrier protein] reductase [EC:1.1.1.100]
ERBA_02342	methanol protein	Methanol-cobalamin methyltransferase B subunit	K04480:methanol---5-hydroxybenzimidazolylcobamide Co-methyltransferase [EC:2.1.1.90]
ERBA_02344	methanol protein	Methanol-cobalamin methyltransferase B subunit	K04480:methanol---5-hydroxybenzimidazolylcobamide Co-methyltransferase [EC:2.1.1.90]
ERBA_01134	Monomethylamine methyltransferase MtmB2	Monomethylamine methyltransferase MtmB Monomethylamine methyltransferase of the archaebacterium Methanosarcina barkeri contains a novel amino acid	K16175:protein scribble
ERBA_01622	Monomethylamine methyltransferase MtmB1	Monomethylamine methyltransferase MtmB Monomethylamine methyltransferase of the archaebacterium	K16175:protein scribble

		Methanosarcina barkeri contains a novel amino acid	
ERBA_01254	Adenine deaminase	Nucleotide transport and metabolism	K01486:adenine deaminase [EC:3.5.4.2]
ERBA_01255	Xanthine phosphoribosyltransferase	Nucleotide transport and metabolism	K00769:xanthine phosphoribosyltransferase [EC:2.4.2.22]
ERBA_02282	Deoxyguanosinetriphosphate triphosphohydrolase-like protein	Nucleotide transport and metabolism	K01129:dGTPase [EC:3.1.5.1]
ERBA_01603	sensory transduction histidine kinase	PAS fold The PAS fold corresponds to the structural domain that has previously been defined as PAS and PAC motifs The PAS fold appears in archaea	
ERBA_00316	Cytochrome c-type biogenesis protein CycK	Posttranslational modification, protein turnover, chaperones	K02926:large subunit ribosomal protein L4
ERBA_00986	ATP-dependent 26S proteasome regulatory subunit Posttranslational modification	Posttranslational modification, protein turnover, chaperones	K13525:transitional endoplasmic reticulum ATPase
ERBA_01018	FKBP-type peptidyl-prolyl cis-trans isomerase 2 Posttranslational modification	Posttranslational modification, protein turnover, chaperones	K02411:flagellar assembly protein FliH
ERBA_01546	Sulfurtransferase TusA	Posttranslational modification, protein turnover, chaperones	K04085:tRNA 2-thiouridine synthesizing protein A [EC:2.8.1.-]
ERBA_01908	Intracellular alkaline protease	Posttranslational modification, protein turnover, chaperones	K01823:isopentenyl-diphosphate delta-isomerase [EC:5.3.3.2]
ERBA_02228	small heat shock protein	Posttranslational modification, protein turnover, chaperones	K13991:photosynthetic reaction center H subunit
ERBA_02229	small heat shock protein	Posttranslational modification, protein turnover, chaperones	K13991:photosynthetic reaction center H subunit
ERBA_01456	pyridoxamine 5'-phosphate oxidase-related FMN-binding protein	Pyridoxamine 5'-phosphate oxidase	
ERBA_00165	transposase	Replication, recombination and repair	
ERBA_00349	UvrABC system protein B	Replication, recombination and repair	K10838:xeroderma pigmentosum group C-complementing protein
ERBA_00517	Transposase	Replication, recombination and repair	
ERBA_00718	putative DNA repair protein	Replication, recombination and repair	
ERBA_00719	DevR family CRISPR-associated autoregulator	Replication, recombination and repair	
ERBA_00721	CRISPR-associated endoribonuclease Cas2	Replication, recombination and repair	K02230:cobaltochelatase CobN [EC:6.6.1.2]
ERBA_00722	CRISPR-associated protein Cas1 1	Replication, recombination and repair	K02434:aspartyl-tRNA(Asn)/glutamyl-tRNA (Gln) amidotransferase subunit B [EC:6.3.5.6 6.3.5.7]
ERBA_00723	CRISPR-associated exonuclease Cas4	Replication, recombination and repair	
ERBA_00724	CRISPR-associated endoribonuclease Cas6 1	Replication, recombination and repair	
ERBA_01039	Putative transposase in snaA-snaB intergenic region	Replication, recombination and repair	
ERBA_01080	Putative transposase in snaA-snaB intergenic region	Replication, recombination and repair	
ERBA_01081	Putative transposase in snaA-snaB intergenic region	Replication, recombination and repair	
ERBA_01285	transposase	Replication, recombination and repair	
ERBA_01327	Putative transposase in snaA-snaB intergenic region	Replication, recombination and repair	K00527:ribonucleoside-triphosphate reductase [EC:1.17.4.2]
ERBA_01388	DNA ligase 1	Replication, recombination and repair	K10744:ribonuclease H2 subunit B
ERBA_01444	Transposase	Replication, recombination and repair	
ERBA_01519	Putative transposase in snaA-snaB intergenic region	Replication, recombination and repair	
ERBA_01520	Putative transposase in snaA-snaB intergenic region	Replication, recombination and repair	
ERBA_01521	putative endonuclease 4	Replication, recombination and repair	K01151:deoxyribonuclease IV [EC:3.1.21.2]
ERBA_01769	Putative transposase in snaA-snaB intergenic region	Replication, recombination and repair	
ERBA_02017	transposase	Replication, recombination and repair	
ERBA_02057	CRISPR-associated protein Cas4-endonuclease Cas1 fusion	Replication, recombination and repair	
ERBA_02277	DNA helicase II - ATP-dependent DNA helicase PcrA	Replication, recombination and repair	K03657:DNA helicase II / ATP-dependent DNA helicase PcrA [EC:3.6.4.12]
ERBA_02286	transposase IS4 family protein	Replication, recombination and repair	
ERBA_02330	transposase	Replication, recombination and repair	
ERBA_01160	SCP-like extracellular	SCP_bacterial: SCP-like extracellular protein domain	K08873:PI-3-kinase-related kinase SMG-1
ERBA_01449	glycine N-methyltransferase	Secondary metabolites biosynthesis, transport and catabolism	K00552:glycine N-methyltransferase [EC:2.1.1.20]
ERBA_00065	Universal stress protein YxiE	Signal transduction mechanisms	
ERBA_00227	sensory transduction histidine kinase	Signal transduction mechanisms	K07716:two-component system, cell cycle sensor histidine kinase PleC [EC:2.7.13.3]
ERBA_00394	PAS-PAC sensor signal transduction histidine kinase	Signal transduction mechanisms	K07679:two-component system, NarL family, sensor histidine kinase EvgS [EC:2.7.13.3]
ERBA_00483	sensory transduction histidine kinase	Signal transduction mechanisms	K02488:two-component system, cell cycle response regulator
ERBA_00829	Phyllosphere-induced regulator PhyR	Signal transduction mechanisms	K07668:two-component system, OmpR family, response regulator VicR
ERBA_00830	Aerobic respiration control sensor protein ArcB	Signal transduction mechanisms	K13924:two-component system, chemotaxis family, CheB/CheR fusion protein [EC:2.1.1.80 3.1.1.61]
ERBA_00984	Universal stress protein	Signal transduction mechanisms	K03412:two-component system, chemotaxis family, response regulator CheB [EC:3.1.1.61]
ERBA_01166	Sporulation kinase C	Signal transduction mechanisms	K07636:two-component system, OmpR family, phosphate regulon sensor histidine kinase PhoR [EC:2.7.13.3]
ERBA_01169	Aerobic respiration control sensor protein ArcB	Signal transduction mechanisms	K11356:two-component system, sensor histidine kinase and response regulator [EC:2.7.13.3]
ERBA_01176	Blue-light-activated histidine kinase	Signal transduction mechanisms	K13924:two-component system, chemotaxis family, CheB/CheR fusion protein [EC:2.1.1.80 3.1.1.61]
ERBA_01213	Virulence sensor protein BvgS	Signal transduction mechanisms	K13924:two-component system, chemotaxis family,

			CheB/CheR fusion protein [EC:2.1.1.80 3.1.1.61]
ERBA_01348	Adaptive-response sensory-kinase SasA	Signal transduction mechanisms	K07636:two-component system, OmpR family, phosphate regulon sensor histidine kinase PhoR [EC:2.7.13.3]
ERBA_01386	C4-dicarboxylate transport sensor protein DctS	Signal transduction mechanisms	K13924:two-component system, chemotaxis family, CheB/CheR fusion protein [EC:2.1.1.80 3.1.1.61]
ERBA_01430	Sensor protein IrlS	Signal transduction mechanisms	K07636:two-component system, OmpR family, phosphate regulon sensor histidine kinase PhoR [EC:2.7.13.3]
ERBA_01433	Aerobic respiration control sensor protein ArcB	Signal transduction mechanisms	K02488:two-component system, cell cycle response regulator
ERBA_01434	PAS-PAC sensor signal transduction histidine kinase	Signal transduction mechanisms	K11356:two-component system, sensor histidine kinase and response regulator [EC:2.7.13.3]
ERBA_01583	Serine-protein kinase RsbW	Signal transduction mechanisms	K13924:two-component system, chemotaxis family, CheB/CheR fusion protein [EC:2.1.1.80 3.1.1.61]
ERBA_01604	Sensor protein LytS	Signal transduction mechanisms	K13924:two-component system, chemotaxis family, CheB/CheR fusion protein [EC:2.1.1.80 3.1.1.61]
ERBA_01675	Sensor protein RcsC	Signal transduction mechanisms	K07679:two-component system, NarL family, sensor histidine kinase EvgS [EC:2.7.13.3]
ERBA_01983	Blue-light-activated histidine kinase	Signal transduction mechanisms	K13924:two-component system, chemotaxis family, CheB/CheR fusion protein [EC:2.1.1.80 3.1.1.61]
ERBA_01282	sulfotransferase	Sulfotransferase family	
ERBA_00966	ArsR family transcriptional regulator	Transcription	
ERBA_01226	Transcription initiation factor IIB	Transcription	K03124:transcription initiation factor TFIIB
ERBA_01293	AbrB family transcriptional regulator	Transcription	
ERBA_01349	ArsR family transcriptional regulator	Transcription	
ERBA_01500	transcriptional regulator	Transcription	K06223:DNA adenine methylase [EC:2.1.1.72]
ERBA_02280	HxlR family transcriptional regulator	Transcription	
ERBA_02149	Translation initiation factor 2 subunit beta	Translation, ribosomal structure and biogenesis	K03238:translation initiation factor 2 subunit 2
ERBA_02281	tRNA-guanine transglycosylase	Translation, ribosomal structure and biogenesis	
ERBA_00077	transposase IS4	Transposase DDE domain Transposase proteins are necessary for efficient DNA transposition	
ERBA_02347	transposase IS4	Transposase DDE domain Transposase proteins are necessary for efficient DNA transposition	
ERBA_00001	hypothetical protein		
ERBA_00025	hypothetical protein		
ERBA_00036	hypothetical protein		
ERBA_00066	Short domain protein		
ERBA_00078	transposase IS4 family protein		
ERBA_00103	protein of unknown function DUF22 protein		
ERBA_00175	hypothetical protein		
ERBA_00176	hypothetical protein		
ERBA_00201	hypothetical protein		
ERBA_00219	hypothetical protein		
ERBA_00244	MTH865-like domain containing protein		
ERBA_00282	hypothetical protein		
ERBA_00283	hypothetical protein		
ERBA_00284	hypothetical protein		
ERBA_00285	hypothetical protein		
ERBA_00286	hypothetical protein		
ERBA_00292	Epoxyqueuosine reductase		K00335:NADH-quinone oxidoreductase subunit F [EC:1.6.5.3]
ERBA_00317	hypothetical protein		
ERBA_00336	hypothetical protein		
ERBA_00348	hypothetical protein		
ERBA_00376	transposase IS4		
ERBA_00377	hypothetical protein		
ERBA_00378	hypothetical protein		
ERBA_00379	hypothetical protein		
ERBA_00385	hypothetical protein		
ERBA_00390	hypothetical protein		
ERBA_00396	hypothetical protein		
ERBA_00426	putative permease protein		
ERBA_00427	putative permease protein		
ERBA_00428	hypothetical protein		
ERBA_00429	NPCBM-associated domain containing protein		
ERBA_00446	hypothetical protein		
ERBA_00484	hypothetical protein		
ERBA_00610	Epoxyqueuosine reductase		K00441:coenzyme F420 hydrogenase beta subunit [EC:1.12.98.1]
ERBA_00629	hypothetical protein		
ERBA_00630	hypothetical protein		
ERBA_00697	glycosyl transferase group 1		
ERBA_00791	hypothetical protein		
ERBA_00808	hypothetical protein		
ERBA_00889	hypothetical protein		
ERBA_00965	B12 binding domain protein		
ERBA_00985	hypothetical protein		
ERBA_00994	hypothetical protein		
ERBA_01000	transposase IS4 family protein		
ERBA_01031	hypothetical protein		
ERBA_01087	hypothetical protein		
ERBA_01088	hypothetical protein		
ERBA_01089	hypothetical protein		

ERBA_01092	Peptidase C39 like protein	
ERBA_01093	hypothetical protein	
ERBA_01094	hypothetical protein	
ERBA_01096	hypothetical protein	
ERBA_01170	hypothetical protein	
ERBA_01183	hypothetical protein	
ERBA_01184	hypothetical protein	
ERBA_01185	hypothetical protein	
ERBA_01202	hypothetical protein	
ERBA_01206	hypothetical protein	
ERBA_01210	hypothetical protein	
ERBA_01214	hypothetical protein	
ERBA_01225	hypothetical protein	
ERBA_01227	hypothetical protein	
ERBA_01229	transposase IS4 family protein	
ERBA_01231	hypothetical protein	
ERBA_01243	hypothetical protein	
ERBA_01276	hypothetical protein	
ERBA_01279	agl cluster protein AglQ	
ERBA_01283	hypothetical protein	
ERBA_01284	hypothetical protein	
ERBA_01289	hypothetical protein	
ERBA_01294	Plasmid pRiA4b ORF-3-like protein	
ERBA_01325	DDE superfamily endonuclease protein	
ERBA_01333	hypothetical protein	
ERBA_01337	hypothetical protein	
ERBA_01338	Serine-threonine-protein kinase pkn1	K02227:adenosylcobinamide-phosphate synthase [EC:6.3.1.10]
ERBA_01339	radical SAM domain-containing protein	
ERBA_01342	hypothetical protein	
ERBA_01345	hypothetical protein	
ERBA_01346	hypothetical protein	
ERBA_01352	hypothetical protein	
ERBA_01365	Matrixin domain containing protein	
ERBA_01366	hypothetical protein	
ERBA_01368	hypothetical protein	
ERBA_01374	transposase IS4 family protein	
ERBA_01396	hypothetical protein	
ERBA_01398	alkylhydroperoxidase like protein	K00627:pyruvate dehydrogenase E2 component (dihydrolipoamide acetyltransferase) [EC:2.3.1.12]
ERBA_01439	hypothetical protein	
ERBA_01440	hypothetical protein	
ERBA_01441	hypothetical protein	
ERBA_01442	hypothetical protein	
ERBA_01445	transposase OrfB family protein	
ERBA_01446	IS605 OrfB family transposase	
ERBA_01459	hypothetical protein	
ERBA_01478	hypothetical protein	
ERBA_01487	transposase IS4 family protein	
ERBA_01488	peptidase S26B signal peptidase	
ERBA_01495	putative ATP-dependent Na+ efflux pump	
ERBA_01499	hypothetical protein	
ERBA_01501	hypothetical protein	
ERBA_01502	UPF0699 transmembrane protein YdbS	
ERBA_01513	flagellin domain-containing protein	
ERBA_01514	hypothetical protein	
ERBA_01517	hypothetical protein	
ERBA_01518	hypothetical protein	
ERBA_01534	hypothetical protein	
ERBA_01537	Reprolysin domain containing protein	
ERBA_01574	hypothetical protein	
ERBA_01577	hypothetical protein	
ERBA_01578	Winged helix-turn-helix protein	
ERBA_01580	hypothetical protein	
ERBA_01581	hypothetical protein	
ERBA_01582	EF hand protein	
ERBA_01584	hypothetical protein	
ERBA_01588	4Fe-4S ferredoxin iron-sulfur binding domain containing protein	K07307:anaerobic dimethyl sulfoxide reductase subunit B (DMSO reductase iron- sulfur subunit)
ERBA_01596	hypothetical protein	
ERBA_01599	hypothetical protein	
ERBA_01623	hypothetical protein	
ERBA_01638	DNA polymerase beta subunit	
ERBA_01670	hypothetical protein	
ERBA_01672	hypothetical protein	
ERBA_01673	transposase IS4	
ERBA_01723	sensory transduction protein kinase	
ERBA_01724	hypothetical protein	
ERBA_01767	hypothetical protein	
ERBA_01774	Phospholipase D-nuclease domain containing protein	
ERBA_01779	DNA alkylation repair enzyme protein	
ERBA_01836	hypothetical protein	
ERBA_01898	hypothetical protein	
ERBA_01899	hypothetical protein	

ERBA_01981	Phosphate-starvation-inducible E protein	
ERBA_01986	hypothetical protein	
ERBA_01987	hypothetical protein	
ERBA_02010	hypothetical protein	
ERBA_02069	hypothetical protein	
ERBA_02092	transposase OrfB family protein	
ERBA_02116	putative cytoplasmic protein	
ERBA_02212	hypothetical protein	
ERBA_02244	hypothetical protein	
ERBA_02245	hypothetical protein	
ERBA_02246	hypothetical protein	
ERBA_02247	hypothetical protein	
ERBA_02248	hypothetical protein	
ERBA_02249	hypothetical protein	
ERBA_02250	hypothetical protein	
ERBA_02251	hypothetical protein	
ERBA_02252	hypothetical protein	
ERBA_02253	Carboxypeptidase regulatory-like domain protein	
ERBA_02254	hypothetical protein	
ERBA_02255	hypothetical protein	
ERBA_02256	hypothetical protein	
ERBA_02257	hypothetical protein	
ERBA_02261	hypothetical protein	
ERBA_02262	hypothetical protein	
ERBA_02263	hypothetical protein	
ERBA_02265	hypothetical protein	
ERBA_02266	hypothetical protein	
ERBA_02268	hypothetical protein	
ERBA_02270	hypothetical protein	
ERBA_02271	hypothetical protein	
ERBA_02273	hypothetical protein	
ERBA_02278	PD-D/EXK nuclease superfamily protein	
ERBA_02279	tRNA-guanine transglycosylase	K11236:cell division control protein 24
ERBA_02283	hypothetical protein	
ERBA_02284	hypothetical protein	
ERBA_02285	hypothetical protein	
ERBA_02287	hypothetical protein	
ERBA_02288	hypothetical protein	
ERBA_02289	hypothetical protein	
ERBA_02290	glycosyl transferase group 1	
ERBA_02291	hypothetical protein	
ERBA_02292	hypothetical protein	
ERBA_02294	hypothetical protein	
ERBA_02295	hypothetical protein	
ERBA_02309	hypothetical protein	
ERBA_02326	hypothetical protein	
ERBA_02346	DsrE/DsrF/DrsH-like domain containing protein	
ERBA_02348	hypothetical protein	

Table S2. Bowtie2 paired-end RNA-seq statistics.

Salinity	5% NaCl			10% NaCl			15% NaCl		
Condition	Sample1	Sample 2	Sample 3	Sample1	Sample 2	Sample 3	Sample1	Sample 2	Sample 3
RNA-seq library	S25	S8	S9	S10	S11	S12	S1	S2	S3
No. of paired-end reads used for mapping	646064	710703	688132	691191	707044	621306	474285	699689	627132

Table S3. COG classification of housekeeping genes in strain ERBA.

Locus ID	Superfamily	COG class	INDIGO annotation
ERBA_00023	COG0065	Amino acid transport and metabolism	3-isopropylmalate dehydratase large subunit 2
ERBA_00906	COG2113	Amino acid transport and metabolism	ABC glycine betaine-L-proline transporter substrate-binding subunit
ERBA_01485	COG0531	Amino acid transport and metabolism	Arginine-ornithine antiporter
ERBA_02190	COG0136	Amino acid transport and metabolism	Aspartate-semialdehyde dehydrogenase
ERBA_02198	COG0040	Amino acid transport and metabolism	ATP phosphoribosyltransferase
ERBA_01436	COG0174	Amino acid transport and metabolism	Glutamine synthetase
ERBA_00740	COG4175	Amino acid transport and metabolism	Glycine betaine transport ATP-binding protein OpuAA
ERBA_00743	COG2113	Amino acid transport and metabolism	Glycine betaine-binding periplasmic protein

ERBA_00742	COG2113	Amino acid transport and metabolism	Glycine betaine-binding periplasmic protein
ERBA_00741	COG4176	Amino acid transport and metabolism	Glycine betaine-L-proline transport system permease protein ProW
ERBA_02196	COG0131	Amino acid transport and metabolism	Imidazoleglycerol-phosphate dehydratase
ERBA_00022	COG0140	Amino acid transport and metabolism	Phosphoribosyl-ATP pyrophosphatase
ERBA_01616	COG0560	Amino acid transport and metabolism	Phosphoserine phosphatase
ERBA_00633	COG0119	Amino acid transport and metabolism	Putative -citramalate synthase CimA
ERBA_00016	COG0119	Amino acid transport and metabolism	putative 2-isopropylmalate synthase
ERBA_00635	COG0440	Amino acid transport and metabolism	putative acetolactate synthase small subunit
ERBA_02333	COG0436	Amino acid transport and metabolism	Putative aminotransferase YugH
ERBA_01667	COG0436	Amino acid transport and metabolism	putative aspartate aminotransferase 1
ERBA_02037	COG0620	Amino acid transport and metabolism	putative methylcobalamin
ERBA_00069	COG0079	Amino acid transport and metabolism	Putative threonine-phosphate decarboxylase
ERBA_01932	COG0112	Amino acid transport and metabolism	Serine hydroxymethyltransferase
ERBA_00868	COG0524	Carbohydrate transport and metabolism	Adenosine kinase
ERBA_00268	COG0057	Carbohydrate transport and metabolism	Glyceraldehyde-3-phosphate dehydrogenase
ERBA_00164	COG0574	Carbohydrate transport and metabolism	Pyruvate phosphate dikinase
ERBA_00853	COG1850	Carbohydrate transport and metabolism	Ribulose bisphosphate carboxylase
ERBA_02085	COG0206	Cell cycle control, cell division, chromosome partitioning	Cell division protein FtsZ 1
ERBA_01566	COG0206	Cell cycle control, cell division, chromosome partitioning	Cell division protein FtsZ 2
ERBA_01043	COG0206	Cell cycle control, cell division, chromosome partitioning	Cell division protein FtsZ 3
ERBA_00505	COG3640	Cell cycle control, cell division, chromosome partitioning	Septum site-determining protein MinD
ERBA_00607	COG1681	Cell motility	Flagellin B2
ERBA_01666	COG0794	Cell wall/membrane/envelope biogenesis	3-hexulose-6-phosphate isomerase
ERBA_01657	COG1361	Cell wall/membrane/envelope biogenesis	hypothetical protein
ERBA_01656	COG1361	Cell wall/membrane/envelope biogenesis	hypothetical protein
ERBA_02098	COG0463	Cell wall/membrane/envelope biogenesis	Protein RfbJ
ERBA_01776	COG1361	Cell wall/membrane/envelope biogenesis	S-layer domain protein
ERBA_01946	COG1004	Cell wall/membrane/envelope biogenesis	UDP-glucose 6-dehydrogenase
ERBA_00871	COG0499	Coenzyme transport and metabolism	Adenosylhomocysteinase
ERBA_01931	COG0190	Coenzyme transport and metabolism	Bifunctional protein FolD
ERBA_01888	COG2082	Coenzyme transport and metabolism	Cobalt-precorrin-8X methylmutase
ERBA_00042	COG0113	Coenzyme transport and metabolism	Delta-aminolevulinic acid dehydratase
ERBA_01930	COG0294	Coenzyme transport and metabolism	Dihydropteroate synthase
ERBA_00049	COG0142	Coenzyme transport and metabolism	Geranylgeranyl diphosphate synthase
ERBA_00041	COG0001	Coenzyme transport and metabolism	Glutamate-1-semialdehyde 21-aminomutase
ERBA_00043	COG0373	Coenzyme transport and metabolism	Glutamyl-tRNA reductase
ERBA_00870	COG2226	Coenzyme transport and metabolism	glycine N-methyltransferase
ERBA_00869	COG2226	Coenzyme transport and metabolism	Malonyl-acyl-carrier protein O-methyltransferase
ERBA_00525	COG3252	Coenzyme transport and metabolism	Methenyltetrahydromethanopterin cyclohydrolase
ERBA_00843	COG4056	Coenzyme transport and metabolism	Methyl-coenzyme M reductase I operon protein C
ERBA_00844	COG4055	Coenzyme transport and metabolism	Methyl-coenzyme M reductase II operon protein D
ERBA_00841	COG4058	Coenzyme transport and metabolism	Methyl-coenzyme M reductase II subunit alpha
ERBA_00845	COG4054	Coenzyme transport and metabolism	Methyl-coenzyme M reductase II subunit beta
ERBA_00842	COG4057	Coenzyme transport and metabolism	Methyl-coenzyme M reductase subunit gamma
ERBA_02113	COG1056	Coenzyme transport and metabolism	Nicotinamide-nucleotide adenylyltransferase
ERBA_02096	COG0422	Coenzyme transport and metabolism	Phosphomethylpyrimidine synthase 2
ERBA_01892	COG2875	Coenzyme transport and metabolism	putative cobalt-precorrin-4 C-methyltransferase
ERBA_01894	COG2242	Coenzyme transport and metabolism	putative cobalt-precorrin-6Y C-methyltransferasedecarboxylating

ERBA_00040	COG0181	Coenzyme transport and metabolism	putative porphobilinogen deaminase
ERBA_01300	COG0214	Coenzyme transport and metabolism	Pyridoxal biosynthesis lyase PdxS
ERBA_00872	COG0192	Coenzyme transport and metabolism	S-adenosylmethionine synthase
ERBA_00044	COG1648	Coenzyme transport and metabolism	Siroheme synthase
ERBA_01807	COG4063	Coenzyme transport and metabolism	Tetrahydromethanopterin S-methyltransferase subunit A
ERBA_01806	COG4062	Coenzyme transport and metabolism	Tetrahydromethanopterin S-methyltransferase subunit B
ERBA_01805	COG4061	Coenzyme transport and metabolism	Tetrahydromethanopterin S-methyltransferase subunit C
ERBA_01804	COG4060	Coenzyme transport and metabolism	Tetrahydromethanopterin S-methyltransferase subunit D
ERBA_01803	COG4059	Coenzyme transport and metabolism	Tetrahydromethanopterin S-methyltransferase subunit E
ERBA_01808	COG4218	Coenzyme transport and metabolism	Tetrahydromethanopterin S-methyltransferase subunit F
ERBA_01809	COG4064	Coenzyme transport and metabolism	Tetrahydromethanopterin S-methyltransferase subunit G
ERBA_01810	COG1962	Coenzyme transport and metabolism	Tetrahydromethanopterin S-methyltransferase subunit H
ERBA_01734n01733MttB	COG5598	Coenzyme transport and metabolism	Trimethylamine methyltransferase MttB
ERBA_01669n01668MttB	COG5598	Coenzyme transport and metabolism	Trimethylamine methyltransferase MttB
ERBA_00143	COG0407	Coenzyme transport and metabolism	Uroporphyrinogen-III decarboxylase Coenzyme metabolism
ERBA_02115	COG2141	Energy production and conversion	510-methylenetetrahydromethanopterin reductase
ERBA_00258	COG1614	Energy production and conversion	Acetyl-CoA decarbonylase-synthase complex subunit beta 1
ERBA_00256	COG2069	Energy production and conversion	Acetyl-CoA decarbonylase-synthase complex subunit delta 2
ERBA_00255	COG1456	Energy production and conversion	Acetyl-CoA decarbonylase-synthase complex subunit gamma 1
ERBA_01393	COG0247	Energy production and conversion	CoB--CoM heterodisulfide reductase 2 iron-sulfursubunit D
ERBA_01615	COG1927	Energy production and conversion	Coenzyme F420-dependent N-methenyltetrahydromethanopterin dehydrogenase Energy production
ERBA_02114	COG1035	Energy production and conversion	Coenzyme F420-reducing hydrogenase
ERBA_00097	COG0644	Energy production and conversion	Digeranylgeranylglycerophospholipid reductase 2
ERBA_01745	COG2878	Energy production and conversion	Electron transport complex protein RnfB
ERBA_01750	COG4656	Energy production and conversion	Electron transport complex protein RnfC
ERBA_01749	COG4658	Energy production and conversion	Electron transport complex protein RnfD
ERBA_00109	COG1005	Energy production and conversion	F oxidoreductase subunit H
ERBA_01634	COG1153	Energy production and conversion	formylmethanofuran dehydrogenase subunit D
ERBA_00224	COG2037	Energy production and conversion	Formylmethanofuran--tetrahydromethanopterin formyltransferase
ERBA_00010	COG0277	Energy production and conversion	Glycolate oxidase subunit GlcD
ERBA_01563	COG3808	Energy production and conversion	K-insensitive pyrophosphate-energized proton pump
ERBA_00112	COG0839	Energy production and conversion	NADH-quinone oxidoreductase chain 6
ERBA_00105	COG0838	Energy production and conversion	NADH-quinone oxidoreductase subunit A
ERBA_00106	COG0377	Energy production and conversion	NADH-quinone oxidoreductase subunit B
ERBA_00107	COG0852	Energy production and conversion	NADH-quinone oxidoreductase subunit B-C-D
ERBA_00108	COG0649	Energy production and conversion	NADH-quinone oxidoreductase subunit D
ERBA_00110	COG1143	Energy production and conversion	NADH-quinone oxidoreductase subunit I
ERBA_00111	COG0839	Energy production and conversion	NADH-quinone oxidoreductase subunit J
ERBA_00113	COG0713	Energy production and conversion	NADH-quinone oxidoreductase subunit K
ERBA_00115	COG1008	Energy production and conversion	NADH-quinone oxidoreductase subunit M
ERBA_00116	COG1007	Energy production and conversion	NADH-quinone oxidoreductase subunit N
ERBA_01636	COG1229	Energy production and conversion	Protein FwdA
ERBA_00675	COG2878	Energy production and conversion	putative ferredoxin
ERBA_02307	COG1592	Energy production and conversion	Putative rubrerythrin
ERBA_01678	COG0674	Energy production and conversion	Pyruvate synthase subunit PorA
ERBA_01677	COG1013	Energy production and conversion	Pyruvate synthase subunit PorB
ERBA_00139	COG2033	Energy production and conversion	Superoxide reductase

ERBA_01635	COG1029	Energy production and conversion	Tungsten-containing formylmethanofuran dehydrogenase 2 subunit B
ERBA_01637	COG2218	Energy production and conversion	Tungsten-containing formylmethanofuran dehydrogenase 2 subunit C
ERBA_00057	COG1155	Energy production and conversion	V-type ATP synthase alpha chain
ERBA_00058	COG1156	Energy production and conversion	V-type ATP synthase beta chain
ERBA_00055	COG1527	Energy production and conversion	V-type ATP synthase subunit C
ERBA_00059	COG1394	Energy production and conversion	V-type ATP synthase subunit D
ERBA_00054	COG1390	Energy production and conversion	V-type ATP synthase subunit E
ERBA_00056	COG1436	Energy production and conversion	V-type ATP synthase subunit F
ERBA_00052	COG1269	Energy production and conversion	V-type ATP synthase subunit I
ERBA_02120	COG0599	Function unknown	alkylhydroperoxidase family protein
ERBA_00953	COG4860	Function unknown	ArsR family transcriptional regulator
ERBA_00475	COG1795	Function unknown	Bifunctional enzyme fae-hp
ERBA_00544	COG1460	Function unknown	DNA-directed RNA polymerase subunit F
ERBA_01557	COG1469	Function unknown	GTP cyclohydrolase MptA
ERBA_00277	COG3430	Function unknown	hypothetical protein
ERBA_01605	COG1698	Function unknown	hypothetical protein
ERBA_01876	COG1844	Function unknown	hypothetical protein
ERBA_01664	COG3365	Function unknown	hypothetical protein
ERBA_00957	COG3390	Function unknown	hypothetical protein
ERBA_00924	COG2450	Function unknown	hypothetical protein
ERBA_00031	COG3388	Function unknown	hypothetical protein
ERBA_01976	COG4089	Function unknown	hypothetical protein
ERBA_00764	COG1422	Function unknown	integral membrane protein
ERBA_01951	COG0316	Function unknown	Iron-sulfur cluster insertion protein ErpA
ERBA_00245	COG1633	Function unknown	methyltransferase
ERBA_01617	COG1340	Function unknown	phosphoserine phosphatase
ERBA_01766	COG1379	Function unknown	PHP-like protein
ERBA_01190	COG1873	Function unknown	PRC-barrel protein
ERBA_02153	COG1873	Function unknown	PRC-barrel protein
ERBA_00592	COG4048	Function unknown	sensory protein
ERBA_01413	COG2138	Function unknown	Sirohydrochlorin cobaltochelatase
ERBA_00223	COG1916	Function unknown	TraB family protein
ERBA_02020	COG3875	Function unknown	transcriptional regulator
ERBA_02118	COG4855	Function unknown	Zinc finger protein
ERBA_01765	COG2047	General function prediction only	3-isopropylmalate dehydratase
ERBA_01741	COG1245	General function prediction only	CoB--CoM heterodisulfide reductase 1 iron-sulfursubunit A
ERBA_00050	COG0535	General function prediction only	Coenzyme PQQ synthesis protein E
ERBA_02136	COG1964	General function prediction only	Cyclic pyranopterin monophosphate synthase
ERBA_01728	COG5012	General function prediction only	Dimethylamine corrinoid protein 2
ERBA_01133	COG5012	General function prediction only	Dimethylamine corrinoid protein 2
ERBA_00626	COG2118	General function prediction only	Double-stranded DNA-binding domain protein
ERBA_00033	COG2249	General function prediction only	FMN-dependent NADH-azoreductase
ERBA_01075	COG1646	General function prediction only	Geranylgeranylglyceryl phosphate synthase
ERBA_01663	COG1100	General function prediction only	GTPase Der
ERBA_00158	COG1355	General function prediction only	Memo-like protein
ERBA_01655	COG1033	General function prediction only	MMPL domain containing protein
ERBA_01658	COG2005	General function prediction only	Molybdenum transport protein ModE
ERBA_01135	COG5012	General function prediction only	Monomethylamine corrinoid protein 1
ERBA_00826	COG0312	General function prediction only	peptidase U62 modulator of DNA gyrase
ERBA_02154	COG0561	General function prediction only	phosphoglycolate phosphatase
ERBA_01975	COG1094	General function prediction only	Polyribonucleotide nucleotidyltransferase
ERBA_00062	COG0714	General function prediction only	Protein CbbQ
ERBA_00864	COG1342	General function prediction only	Protein of unknown function DUF134
ERBA_00024	COG0618	General function prediction only	putative manganese-dependent inorganic pyrophosphatase
ERBA_00047	COG0535	General function prediction only	Putative mycofactocin radical SAM maturase MftC
ERBA_01302	COG3576	General function prediction only	Pyridoxine-pyridoxamine 5'-phosphate oxidase
ERBA_01640	COG3291	General function prediction only	S-layer-related protein
ERBA_00037	COG5485	General function prediction only	SnoaL-like polyketide cyclase protein
ERBA_00925	COG1779	General function prediction only	thioredoxin
ERBA_02149	COG3269	General function prediction only	Translation initiation factor 2 subunit beta
ERBA_02150	COG3269	General function prediction only	Translation initiation factor 2 subunit beta
ERBA_00060	COG3269	General function prediction only	Translation initiation factor 2 subunit beta
ERBA_01732	COG5012	General function prediction only	Trimethylamine corrinoid protein 2
ERBA_02315	COG1350	General function prediction only	Tryptophan synthase beta chain 2
ERBA_00459	COG3364	General function prediction only	Zn-ribbon containing protein
ERBA_01661	COG0725	Inorganic ion transport and metabolism	Molybdate-tungstate-binding protein WtpA
ERBA_01979	COG0704	Inorganic ion transport and metabolism	Phosphate-specific transport system accessory protein PhoU
ERBA_00018	COG0573	Inorganic ion transport and metabolism	putative ABC transporter permease protein YqgH
ERBA_02086	COG2443	Intracellular trafficking, secretion, and vesicular transport	Protein translocase subunit SecE
ERBA_00766	COG0201	Intracellular trafficking, secretion, and vesicular transport	Protein translocase subunit SecY
ERBA_00620	COG0552	Intracellular trafficking, secretion, and vesicular transport	Signal recognition particle receptor FtsY
ERBA_02126	COG1260	Lipid transport and metabolism	myo-inositol-1-phosphate synthase
ERBA_01916	COG0473	Multiple classes	3-isopropylmalate dehydrogenase
ERBA_00454	COG0329	Multiple classes	4-hydroxy-tetrahydrodipicolinate synthase
ERBA_00061	COG0513	Multiple classes	DEAD-box ATP-dependent RNA helicase CshA
ERBA_00557	COG0145	Multiple classes	Hydantoin utilization protein A
ERBA_00636	COG0059	Multiple classes	Ketol-acid reductoisomerase
ERBA_00114	COG1009	Multiple classes	NADH-quinone oxidoreductase subunit L
ERBA_02298	COG1474	Multiple classes	ORC1-type DNA replication protein 2
ERBA_01432	COG0745	Multiple classes	Response regulator MprA

ERBA_01974	COG1718	Multiple classes	RIO-type serine-threonine-protein kinase Rio1
ERBA_01691	COG0591	Multiple classes	Sodium-proline symporter
ERBA_00765	COG0563	Nucleotide transport and metabolism	Adenylate kinase
ERBA_01901	COG1328	Nucleotide transport and metabolism	Anaerobic ribonucleoside-triphosphate reductase
ERBA_02054	COG0138	Nucleotide transport and metabolism	Bifunctional purine biosynthesis protein PurH
ERBA_00039	COG0167	Nucleotide transport and metabolism	Dihydroorotate dehydrogenase B catalytic subunit
ERBA_00552	COG0041	Nucleotide transport and metabolism	N5-carboxyaminoimidazole ribonucleotide mutase
ERBA_01305	COG0105	Nucleotide transport and metabolism	Nucleoside diphosphate kinase
ERBA_02144	COG0046	Nucleotide transport and metabolism	Phosphoribosylformylglycinamidine synthase 2
ERBA_02319	COG0034	Nucleotide transport and metabolism	putative amidophosphoribosyltransferase
ERBA_01939	COG0856	Nucleotide transport and metabolism	PyrE-like protein 1
ERBA_02161	COG0459	Posttranslational modification, protein turnover, chaperones	60 kDa chaperonin
ERBA_00825	COG1067	Posttranslational modification, protein turnover, chaperones	Archaeal Lon protease
ERBA_01925	COG3118	Posttranslational modification, protein turnover, chaperones	Bifunctional thioredoxin reductase-thioredoxin
ERBA_00092	COG0484	Posttranslational modification, protein turnover, chaperones	Chaperone protein DnaJ
ERBA_00091	COG0443	Posttranslational modification, protein turnover, chaperones	Chaperone protein DnaK
ERBA_00048	COG2332	Posttranslational modification, protein turnover, chaperones	Cytochrome c-type biogenesis protein CcmE
ERBA_02041	COG1047	Posttranslational modification, protein turnover, chaperones	FKBP-type peptidyl-prolyl cis-trans isomerase 2
			Posttranslational modification
ERBA_00621	COG1730	Posttranslational modification, protein turnover, chaperones	Prefoldin subunit alpha
ERBA_00412	COG1382	Posttranslational modification, protein turnover, chaperones	Prefoldin subunit beta
ERBA_00837	COG0501	Posttranslational modification, protein turnover, chaperones	Protease HtpX
ERBA_00420	COG0638	Posttranslational modification, protein turnover, chaperones	Proteasome subunit alpha
ERBA_00248	COG0638	Posttranslational modification, protein turnover, chaperones	Proteasome subunit beta
ERBA_01873	COG1222	Posttranslational modification, protein turnover, chaperones	Proteasome-activating nucleotidase
ERBA_02152	COG0464	Posttranslational modification, protein turnover, chaperones	Protein CdcH
ERBA_00090	COG0576	Posttranslational modification, protein turnover, chaperones	Protein GrpE
ERBA_00863	COG0071	Posttranslational modification, protein turnover, chaperones	Small heat shock protein C4
ERBA_01706	COG0459	Posttranslational modification, protein turnover, chaperones	Thermosome subunit
ERBA_02299	COG0459	Posttranslational modification, protein turnover, chaperones	Thermosome subunit beta
ERBA_01469	COG0459	Posttranslational modification, protein turnover, chaperones	Thermosome subunit beta
ERBA_01199	COG0592	Replication, recombination and repair	DNA polymerase sliding clamp
ERBA_00006	COG0358	Replication, recombination and repair	DNA primase
ERBA_02155	COG0468	Replication, recombination and repair	DNA repair
ERBA_00958	COG1599	Replication, recombination and repair	replication factor A
ERBA_01849	COG0675	Replication, recombination and repair	Transposase
ERBA_00032	COG0179	Secondary metabolites biosynthesis, transport and catabolism	2-hydroxyhepta-24-diene-17-dioate isomerase
ERBA_02035	COG0467	Signal transduction mechanisms	Circadian clock protein kinase KaiC
ERBA_02336	COG0467	Signal transduction mechanisms	Circadian clock protein kinase KaiC
ERBA_01977	COG0467	Signal transduction mechanisms	Circadian clock protein kinase KaiC
ERBA_02031	COG3448	Signal transduction mechanisms	Inosine-5'-monophosphate dehydrogenase
ERBA_02030	COG3448	Signal transduction mechanisms	Inosine-5'-monophosphate dehydrogenase
ERBA_01641	COG0467	Signal transduction mechanisms	KaiC
ERBA_01484	COG0589	Signal transduction mechanisms	Universal stress protein
ERBA_00686	COG0086	Transcription	DNA-directed RNA polymerase subunit A'
ERBA_00685	COG0086	Transcription	DNA-directed RNA polymerase subunit A"
ERBA_00687	COG0085	Transcription	DNA-directed RNA polymerase subunit B'
ERBA_00688	COG0085	Transcription	DNA-directed RNA polymerase subunit B"
ERBA_00753	COG0202	Transcription	DNA-directed RNA polymerase subunit D
ERBA_00689	COG2012	Transcription	DNA-directed RNA polymerase subunit H
ERBA_00150	COG1644	Transcription	DNA-directed RNA polymerase subunit N
ERBA_01934	COG1395	Transcription	Helix-turn-helix domain containing protein
ERBA_02332	COG1522	Transcription	HTH-type transcriptional regulator LrpA
ERBA_00021	COG2524	Transcription	Inosine-5'-monophosphate dehydrogenase
ERBA_01778	COG1510	Transcription	MarR family transcriptional regulator
ERBA_01952	COG1308	Transcription	Nascent polypeptide-associated complex protein
ERBA_00046	COG1522	Transcription	Protein NirD
ERBA_00045	COG1522	Transcription	Protein NirD
ERBA_00915	COG1958	Transcription	Putative snRNP Sm-like protein
ERBA_00683	COG0195	Transcription	putative transcription termination protein NusA
ERBA_01423	COG1095	Transcription	Ribonuclease R

ERBA_02099	COG2101	Transcription	TATA-box-binding protein 1
ERBA_02087	COG0250	Transcription	Transcription elongation factor Spt5
ERBA_00167	COG1405	Transcription	Transcription initiation factor IIB
ERBA_01571	COG3432	Transcription	Winged helix-turn-helix protein
ERBA_00678	COG0051	Translation, ribosomal structure and biogenesis	30S ribosomal protein S10
ERBA_00754	COG0100	Translation, ribosomal structure and biogenesis	30S ribosomal protein S11
ERBA_00682	COG0048	Translation, ribosomal structure and biogenesis	30S ribosomal protein S12
ERBA_00756	COG0099	Translation, ribosomal structure and biogenesis	30S ribosomal protein S13
ERBA_00071	COG0184	Translation, ribosomal structure and biogenesis	30S ribosomal protein S15-S13e
ERBA_00455	COG1383	Translation, ribosomal structure and biogenesis	30S ribosomal protein S17e
ERBA_00779	COG0186	Translation, ribosomal structure and biogenesis	30S ribosomal protein S17P
ERBA_00784	COG0185	Translation, ribosomal structure and biogenesis	30S ribosomal protein S19
ERBA_00627	COG2238	Translation, ribosomal structure and biogenesis	30S ribosomal protein S19e
ERBA_00157	COG0052	Translation, ribosomal structure and biogenesis	30S ribosomal protein S2
ERBA_01426	COG2004	Translation, ribosomal structure and biogenesis	30S ribosomal protein S24e
ERBA_01763	COG2051	Translation, ribosomal structure and biogenesis	30S ribosomal protein S27e
ERBA_01307	COG2053	Translation, ribosomal structure and biogenesis	30S ribosomal protein S28e
ERBA_00782	COG0092	Translation, ribosomal structure and biogenesis	30S ribosomal protein S3
ERBA_00027	COG1890	Translation, ribosomal structure and biogenesis	30S ribosomal protein S3Ae
ERBA_00755	COG0522	Translation, ribosomal structure and biogenesis	30S ribosomal protein S4
ERBA_00776	COG1471	Translation, ribosomal structure and biogenesis	30S ribosomal protein S4e
ERBA_00769	COG0098	Translation, ribosomal structure and biogenesis	30S ribosomal protein S5
ERBA_01303	COG2125	Translation, ribosomal structure and biogenesis	30S ribosomal protein S6e
ERBA_00681	COG0049	Translation, ribosomal structure and biogenesis	30S ribosomal protein S7
ERBA_00774	COG0096	Translation, ribosomal structure and biogenesis	30S ribosomal protein S8
ERBA_02331	COG2007	Translation, ribosomal structure and biogenesis	30S ribosomal protein S8e
ERBA_00149	COG0103	Translation, ribosomal structure and biogenesis	30S ribosomal protein S9
ERBA_02089	COG0081	Translation, ribosomal structure and biogenesis	50S ribosomal protein L1
ERBA_02090	COG0244	Translation, ribosomal structure and biogenesis	50S ribosomal protein L10
ERBA_00300	COG0197	Translation, ribosomal structure and biogenesis	50S ribosomal protein L10e
ERBA_02088	COG0080	Translation, ribosomal structure and biogenesis	50S ribosomal protein L11
ERBA_02091	COG2058	Translation, ribosomal structure and biogenesis	50S ribosomal protein L12
ERBA_00148	COG0102	Translation, ribosomal structure and biogenesis	50S ribosomal protein L13P
ERBA_00778	COG0093	Translation, ribosomal structure and biogenesis	50S ribosomal protein L14P
ERBA_00424	COG1632	Translation, ribosomal structure and biogenesis	50S ribosomal protein L15e
ERBA_00767	COG0200	Translation, ribosomal structure and biogenesis	50S ribosomal protein L15P
ERBA_00147	COG1727	Translation, ribosomal structure and biogenesis	50S ribosomal protein L18e
ERBA_00770	COG0256	Translation, ribosomal structure and biogenesis	50S ribosomal protein L18P
ERBA_00771	COG2147	Translation, ribosomal structure and biogenesis	50S ribosomal protein L19e
ERBA_00785	COG0090	Translation, ribosomal structure and biogenesis	50S ribosomal protein L2
ERBA_00783	COG0091	Translation, ribosomal structure and biogenesis	50S ribosomal protein L22P
ERBA_00786	COG0089	Translation, ribosomal structure and biogenesis	50S ribosomal protein L23P
ERBA_01306	COG2075	Translation, ribosomal structure and biogenesis	50S ribosomal protein L24e
ERBA_00788	COG0087	Translation, ribosomal structure and biogenesis	50S ribosomal protein L3
ERBA_00684	COG1911	Translation, ribosomal structure and biogenesis	50S ribosomal protein L30e
ERBA_00768	COG1841	Translation, ribosomal structure and biogenesis	50S ribosomal protein L30P
ERBA_00624	COG2097	Translation, ribosomal structure and biogenesis	50S ribosomal protein L31e
ERBA_00772	COG1717	Translation, ribosomal structure and biogenesis	50S ribosomal protein L32e
ERBA_00415	COG1997	Translation, ribosomal structure and biogenesis	50S ribosomal protein L37Ae
ERBA_00787	COG0088	Translation, ribosomal structure	50S ribosomal protein L4

		and biogenesis	
ERBA_01074	COG1552	Translation, ribosomal structure and biogenesis	50S ribosomal protein L40e
ERBA_01762	COG1631	Translation, ribosomal structure and biogenesis	50S ribosomal protein L44e
ERBA_00775	COG0094	Translation, ribosomal structure and biogenesis	50S ribosomal protein L5
ERBA_00773	COG0097	Translation, ribosomal structure and biogenesis	50S ribosomal protein L6
ERBA_01308	COG1358	Translation, ribosomal structure and biogenesis	50S ribosomal protein L7Ae
ERBA_00625	COG2117	Translation, ribosomal structure and biogenesis	7-cyano-7-deazaguanine synthase
ERBA_00679	COG5256	Translation, ribosomal structure and biogenesis	Elongation factor 1-alpha
ERBA_00680	COG0480	Translation, ribosomal structure and biogenesis	Elongation factor 2
ERBA_00166	COG3277	Translation, ribosomal structure and biogenesis	H-ACA RNA-protein complex component Gar1
ERBA_00545	COG1491	Translation, ribosomal structure and biogenesis	hypothetical protein
ERBA_00922	COG0060	Translation, ribosomal structure and biogenesis	Isoleucine--tRNA ligase
ERBA_02063	COG0143	Translation, ribosomal structure and biogenesis	Methionine--tRNA ligase
ERBA_01924	COG1236	Translation, ribosomal structure and biogenesis	mRNA 3-end processing factor
ERBA_00252	COG1499	Translation, ribosomal structure and biogenesis	NMD3 family protein
ERBA_02142	COG0023	Translation, ribosomal structure and biogenesis	Protein translation factor SUI1
ERBA_00417	COG0689	Translation, ribosomal structure and biogenesis	putative exosome complex exonuclease 1
ERBA_00416	COG2123	Translation, ribosomal structure and biogenesis	putative exosome complex exonuclease 2
ERBA_00418	COG1097	Translation, ribosomal structure and biogenesis	putative exosome complex RNA-binding protein 1
ERBA_00002	COG1859	Translation, ribosomal structure and biogenesis	putative RNA 2'-phosphotransferase
ERBA_01304	COG0532	Translation, ribosomal structure and biogenesis	putative translation initiation factor IF-2
ERBA_00780	COG1588	Translation, ribosomal structure and biogenesis	Ribonuclease P protein component 1
ERBA_00422	COG1603	Translation, ribosomal structure and biogenesis	Ribonuclease P protein component 3
ERBA_01652	COG2023	Translation, ribosomal structure and biogenesis	Ribonuclease P protein component 4
ERBA_00419	COG1500	Translation, ribosomal structure and biogenesis	Ribosome maturation protein SDO1
ERBA_00423	COG1325	Translation, ribosomal structure and biogenesis	RNA binding protein
ERBA_01973	COG0361	Translation, ribosomal structure and biogenesis	Translation initiation factor 1A 1
ERBA_01597	COG0361	Translation, ribosomal structure and biogenesis	Translation initiation factor 1A 2
ERBA_01764	COG1093	Translation, ribosomal structure and biogenesis	Translation initiation factor 2 subunit alpha
ERBA_00301	COG1601	Translation, ribosomal structure and biogenesis	Translation initiation factor 2 subunit beta
ERBA_00200	COG0231	Translation, ribosomal structure and biogenesis	Translation initiation factor 5A
ERBA_00623	COG1976	Translation, ribosomal structure and biogenesis	Translation initiation factor 6
ERBA_00026	COG0172	Translation, ribosomal structure and biogenesis	Type-2 serine--tRNA ligase
ERBA_00777	-	-	50S ribosomal protein L24P
ERBA_00781	-	-	50S ribosomal protein L29P
ERBA_00529	-	-	ArsR family transcriptional regulator
ERBA_00914	-	-	B-block binding subunit of TFIIIC protein
ERBA_00534	-	-	cell surface lipoprotein
ERBA_01556	-	-	chromosomal protein
ERBA_01392	-	-	CoB--CoM heterodisulfide reductase 2 subunit E
ERBA_01751	-	-	cytochrome C
ERBA_00960n00961MtbB	-	-	Dimethylamine methyltransferase MtbB3
ERBA_01638	-	-	DNA polymerase beta subunit
ERBA_00117	-	-	F dehydrogenase subunit FpoO
ERBA_00051	-	-	H-transporting ATP synthase subunit H
ERBA_00130	-	-	hypothetical protein
ERBA_01730	-	-	hypothetical protein
ERBA_01729	-	-	hypothetical protein
ERBA_00234	-	-	hypothetical protein
ERBA_00276	-	-	hypothetical protein
ERBA_00930	-	-	hypothetical protein
ERBA_00916	-	-	hypothetical protein
ERBA_01665	-	-	hypothetical protein
ERBA_01829	-	-	hypothetical protein
ERBA_01752	-	-	hypothetical protein
ERBA_02105	-	-	hypothetical protein
ERBA_00731	-	-	hypothetical protein
ERBA_01581	-	-	hypothetical protein
ERBA_01988	-	-	hypothetical protein
ERBA_01580	-	-	hypothetical protein
ERBA_00555	-	-	hypothetical protein
ERBA_00882	-	-	methanol protein

ERBA_01991n01992MtmB	-	-	Monomethylamine methyltransferase MtmB
ERBA_01134	-	-	Monomethylamine methyltransferase MtmB2
ERBA_01906	-	-	Preprotein translocase subunit SecG
ERBA_02055	-	-	Putative adenylate cyclase 3
ERBA_00053	-	-	putative ATPase proteolipid chain
ERBA_00561	-	-	small heat shock protein
ERBA_02347	-	-	transposase IS4
ERBA_01731	-	-	trimethylamine permease

Table S4. 256 up-regulated genes at 15% salinity relative to 10% salinity. (Genes with log2 fold change greater than 1 are in bold.)

Gene ID	INDIGO annotation	log2 fold change	COG classification	FPKM avg. 5% NaCl	FPKM avg. 10% NaCl	FPKM avg. 15% NaCl	std.dev for 5% NaCl	std.dev for 10% NaCl	std.dev for 15% NaCl
ERBA_00486	**dienelactone hydrolase**	**1.72**	**Amino acid transport and metabolism**	**409.13**	**59.81**	**261.9**	**207.83**	**12.3**	**43.42**
ERBA_01555	**Acetylglutamate kinase**	**1.68**	**Amino acid transport and metabolism**	**159.88**	**49.27**	**211.97**	**32.13**	**11.05**	**48.4**
ERBA_01116	**N-acetylglutamate synthase**	**1.53**	**Amino acid transport and metabolism**	**201.07**	**96.18**	**369.59**	**42.27**	**9.5**	**41.6**
ERBA_02050	**L-tyrosine decarboxylase**	**1.33**	**Amino acid transport and metabolism**	**862.86**	**146.71**	**488.3**	**266.3**	**14.41**	**56.17**
ERBA_01624	**D-serine-D-alanine-glycine transporter**	**1.23**	**Amino acid transport and metabolism**	**129.65**	**46.39**	**145.28**	**19.21**	**11.12**	**22.93**
ERBA_02242	**putative cysteine desulfurase**	**1.03**	**Amino acid transport and metabolism**	**2073.18**	**180.67**	**488.52**	**1009.17**	**12.94**	**77.68**
ERBA_01098	**NADPH-dependent glutamate synthase large subunit**	**1**	**Amino acid transport and metabolism**	**136.98**	**44.32**	**117.33**	**41.15**	**10.05**	**33.36**
ERBA_01381	Cystathionine gamma-synthase	0.98	Amino acid transport and metabolism	219.03	77.45	202.33	51.27	15.9	21.47
ERBA_00315	O-acetylserine sulfhydrylase	0.88	Amino acid transport and metabolism	65.72	26.95	66.87	12.75	2.1	12.74
ERBA_01875	Pyrroline-5-carboxylate reductase	0.85	Amino acid transport and metabolism	512.85	318.97	764.96	60.78	18.36	70.99
ERBA_00501	LL-diaminopimelate aminotransferase	0.8	Amino acid transport and metabolism	218.53	146.35	337.2	40.55	27.31	31.22
ERBA_02215	peptidase M24	0.8	Amino acid transport and metabolism	128.86	99.27	227.79	18.7	3.71	8.94
ERBA_00945	Cysteate synthase	0.74	Amino acid transport and metabolism	381.27	210.22	462.55	21.66	17.45	18.89
ERBA_01841	Acetylornithine aminotransferase	0.71	Amino acid transport and metabolism	318.21	120.05	261.62	31.23	4.94	38.72
ERBA_01030	3-isopropylmalate dehydratase large subunit 1	0.65	Amino acid transport and metabolism	306.18	130.43	270.31	80.08	16.44	40.72
ERBA_02037	putative methylcobalamin	0.61	Amino acid transport and metabolism	448.46	351.5	710.73	60.54	20.62	99.26
ERBA_00264	Putative homocitrate synthase AksA	0.58	Amino acid transport and metabolism	286.25	206.74	409.26	24.7	14.31	29.69
ERBA_00813	Ornithine carbamoyltransferase	0.57	Amino acid transport and metabolism	266.37	205.92	403.84	19.48	20.15	14.64
ERBA_00793	Glutamate 5-kinase	0.57	Amino acid transport and metabolism	147.01	145.49	285.59	19.28	11.65	10.22
ERBA_00498	Argininosuccinate lyase	0.55	Amino acid transport and metabolism	215.99	110.89	216.98	9.56	9.88	40.08
ERBA_00740	Glycine betaine transport ATP-binding protein OpuAA	0.55	Amino acid transport and metabolism	1025.99	772.58	1491.22	102.69	39.27	30.6
ERBA_01814	Aspartokinase	0.53	Amino acid transport and metabolism	178.82	178.38	339.99	24.63	9.12	11
ERBA_01099	**CoB--CoM heterodisulfide reductase 1 iron-sulfursubunit A**	**1.1**	**Amino acid transport and metabolism /Energy production and conversion**	**134.25**	**46.61**	**131.83**	**48.28**	**14.33**	**23.62**
ERBA_01382	Homoserine O-acetyltransferase	0.82	Amino acid transport and metabolism /General function prediction only	171.29	80.48	187	40.8	7.13	7.04
ERBA_00431	**cupin**	**2.49**	**Carbohydrate transport and metabolism**	**405.05**	**59.45**	**446.87**	**64.5**	**10.75**	**58.36**
ERBA_00119	**Fructose-16-bisphosphatase FBPase V**	**1.82**	**Carbohydrate transport and metabolism**	**78.8**	**31.55**	**149.27**	**5.11**	**1.94**	**28.59**
ERBA_01117	**glucan 14-alpha-glucosidase**	**1.14**	**Carbohydrate transport and metabolism**	**119.73**	**59.93**	**174.77**	**28.59**	**4.2**	**8.63**
ERBA_01683	Protein FrlC	0.88	Carbohydrate transport and metabolism	261.19	133.37	329.22	42.65	39.03	44.67
ERBA_00035	glycoside hydrolase 15-like protein	0.83	Carbohydrate transport and metabolism	964.5	167.69	394.65	333.62	21.79	88.08
ERBA_01253	Putative alpha-amylase	0.74	Carbohydrate transport and metabolism	135.47	101.21	223.98	7.99	12.95	20.86
ERBA_00474	Triosephosphate isomerase	0.7	Carbohydrate transport and metabolism	384.06	224.58	482.58	28.82	19.03	58.25
ERBA_00221	23-bisphosphoglycerate-independent phosphoglycerate mutase 1	0.67	Carbohydrate transport and metabolism	271.86	130.07	273.69	23.44	9.31	15.49
ERBA_00632	Cellulase M	0.63	Carbohydrate transport and metabolism	348.52	245.61	502.3	80.35	7.92	23.21
ERBA_00434	**PP-loop domain containing protein**	**1.4**	**Cell cycle control, cell division, chromosome partitioning**	**725.07**	**89.16**	**311.31**	**127.18**	**3.32**	**25.08**
ERBA_01860	UDP-4-amino-4-deoxy-L-arabinose--oxoglutarate aminotransferase	0.98	Cell wall/membrane/envelope biogenesis	147.89	58.8	154.17	53.98	9.41	6.6
ERBA_01403	**2--3'-dephosphocoenzyme-A synthase**	**1.17**	**Coenzyme transport and metabolism**	**190.04**	**77.75**	**232.93**	**6.56**	**3.82**	**21.71**
ERBA_00562	**Molybdopterin synthase catalytic subunit**	**1.16**	**Coenzyme transport and metabolism**	**74.86**	**52.62**	**156.14**	**16.82**	**5.55**	**18.69**
ERBA_00525	**Methenyltetrahydromethanopterin cyclohydrolase**	**1**	**Coenzyme transport and metabolism**	**968.8**	**609.83**	**1615.31**	**343.46**	**17.99**	**103.37**
ERBA_02052	prenyltransferase	0.96	Coenzyme transport and metabolism	119.4	36.29	94.69	23.36	1.98	17.54
ERBA_01335	**phenylacetate--CoA ligase**	**0.91**	**Coenzyme transport and metabolism**	**60.98**	**19.64**	**49.13**	**9.25**	**2.68**	**6.21**
ERBA_01201	Cobyrinic acid AC-diamide synthase	0.88	Coenzyme transport and metabolism	445.47	207.56	503.12	37.05	13.1	51.69
ERBA_02166	putative cobalt-precorrin-6Y C-methyltransferasedecarboxylating	0.81	Coenzyme transport and metabolism	164.1	108.9	251.33	40.35	15.19	8.2
ERBA_00074	**Biotin synthase**	**0.8**	**Coenzyme transport and metabolism**	**285.42**	**73.75**	**170.89**	**66.84**	**10.21**	**20.63**
ERBA_00877	Nicotinate phosphoribosyltransferase	0.69	Coenzyme transport and metabolism	91.86	64.14	135.96	18.71	11.84	20.22
ERBA_01300	Pyridoxal biosynthesis lyase PdxS	0.6	Coenzyme transport and metabolism	574.45	360.19	722.53	106.71	7.33	7.59
ERBA_00642	Phosphomethylpyrimidine synthase 1	0.57	Coenzyme transport and metabolism	138.63	82.88	162.53	8.52	7.86	1.61
ERBA_02316	Hydroxymethylpyrimidine-phosphomethylpyrimidine kinase	0.77	Coenzyme transport and metabolism /Function unknown	196.65	105.85	239.07	19.03	11.01	19.89
ERBA_01920	**Multidrug export protein MepA**	**1.16**	**Defense mechanisms**	**160.48**	**41.65**	**123.23**	**32.14**	**12.47**	**22.96**

ERBA_00654	Nod factor export ATP-binding protein I	0.8	Defense mechanisms	156.19	59.06	137.44	21.52	3.58	21.93
ERBA_00432	**Hydroxylamine reductase**	**3**	**Energy production and conversion**	**1328.38**	**176.1**	**1862.17**	**266.65**	**18.93**	**268.65**
ERBA_00577	**Putative fumarate hydratase subunit beta**	**2.14**	**Energy production and conversion**	**212.48**	**47.65**	**280.68**	**65.04**	**12.96**	**67.58**
ERBA_00576	**Putative fumarate hydratase subunit alpha**	**1.6**	**Energy production and conversion**	**343.14**	**79.47**	**320.21**	**78.79**	**8.87**	**43.32**
ERBA_00903	**Tungsten-containing aldehyde ferredoxin oxidoreductase**	**1.4**	**Energy production and conversion**	**113.18**	**25.72**	**90.2**	**39.94**	**5.93**	**23.99**
ERBA_00260	**Acetyl-CoA decarbonylase-synthase complex subunit alpha**	**1.18**	**Energy production and conversion**	**854.04**	**263.74**	**793.28**	**468.44**	**26.04**	**264.92**
ERBA_00292	**Epoxyqueuosine reductase**	**1.16**	**Energy production and conversion**	**108.38**	**46.22**	**137.59**	**48.36**	**8.98**	**13.88**
ERBA_01627	**Putative superoxide reductase**	**1.13**	**Energy production and conversion**	**108.88**	**36.28**	**105.61**	**48.62**	**7.69**	**13.15**
ERBA_01472	**NADPH dehydrogenase**	**1.01**	**Energy production and conversion**	**85.58**	**53.98**	**143.21**	**17.53**	**7.24**	**7.77**
ERBA_01014	adenosine tRNA methylthiotransferase MiaB	0.96	Energy production and conversion	366.31	107.29	276.13	120.4	10.52	29.16
ERBA_00259	Acetyl-CoA decarbonylase-synthase complex subunit epsilon	0.95	Energy production and conversion	654.94	227.13	581.78	357.2	38.1	179.98
ERBA_01103	CoB--CoM heterodisulfide reductase 1 subunit B	0.81	Energy production and conversion	576.5	126.38	294.25	205.02	9.47	27.61
ERBA_01132	CoB--CoM heterodisulfide reductase 1 iron-sulfursubunit A	0.81	Energy production and conversion	236.27	82.39	190.55	83.73	3.78	28.12
ERBA_01191	Malate dehydrogenase	0.78	Energy production and conversion	344.42	277.59	628.75	44.76	16.82	54.6
ERBA_00202	Tungsten-containing aldehyde ferredoxin oxidoreductase	0.74	Energy production and conversion	108.81	40.62	89.14	20.02	1.72	7.82
ERBA_00400	Indolepyruvate oxidoreductase subunit IorA	0.73	Energy production and conversion	167.28	108.45	238.37	39.85	14.25	27.57
ERBA_01482	Acetyl-CoA decarbonylase-synthase complex subunit gamma	0.7	Energy production and conversion	333.07	140.66	299.5	63.21	18.36	28.21
ERBA_01563	K-insensitive pyrophosphate-energized proton pump	0.6	Energy production and conversion	662.63	640.14	1280.9	83.18	25.57	154.52
ERBA_01481	Tungsten-containing formylmethanofuran dehydrogenase 2 subunit B	0.55	Energy production and conversion	373.03	154.77	298.49	80.93	8	24.61
ERBA_00357	Oxaloacetate decarboxylase alpha chain	0.63	Energy production and conversion /Lipid transport and metabolism	331.89	117.77	240.12	60.35	4.7	56.52
ERBA_00952	**Pyruvoyl-dependent arginine decarboxylase 1**	**1.37**	**Function unknown**	**96.65**	**41.8**	**141.46**	**15.47**	**11.97**	**33.33**
ERBA_00674	**transporter**	**1.29**	**Function unknown**	**141.42**	**29.69**	**97.84**	**45.5**	**8.12**	**16.47**
ERBA_01502	**UPF0699 transmembrane protein YdbS (contains bPH2 (bacterial pleckstrin homology) domain)**	**1.28**	**Function unknown**	**86.41**	**39.9**	**130.09**	**25.02**	**12.48**	**22.47**
ERBA_00526	**hypothetical protein**	**1.27**	**Function unknown**	**391.95**	**179.09**	**577.82**	**16.55**	**37.66**	**110.6**
ERBA_00578	**hypothetical protein**	**1.24**	**Function unknown**	**191.38**	**56.95**	**178.01**	**54.01**	**3.75**	**16.25**
ERBA_01470	**hypothetical protein**	**1.16**	**Function unknown**	**82.72**	**27.64**	**83.1**	**28.41**	**9.61**	**1.96**
ERBA_01514	**hypothetical protein (archaeal flagellin N-terminal-like domain)**	**1.1**	**Function unknown**	**57.25**	**41.96**	**119.95**	**10.62**	**8.79**	**18.73**
ERBA_00588	**putative containing protein**	**1.03**	**Function unknown**	**182.73**	**67.11**	**181.67**	**45.93**	**8.83**	**25.15**
ERBA_02192	**Adenosylcobinamide amidohydrolase**	**1.02**	**Function unknown**	**165.74**	**74.61**	**200.46**	**41.14**	**15.79**	**17.24**
ERBA_02127	rubrerythrin	0.98	Function unknown	288.05	99.77	262.16	12.72	24.59	16.04
ERBA_00326	2-phospho-L-lactate transferase	0.97	Function unknown	100.16	45.85	120.6	11.01	12.28	34.57
ERBA_00381	Phosphoenolpyruvate carboxylase	0.97	Function unknown	122.4	35	91.69	26.99	3.02	15.73
ERBA_00541	tRNA -methyltransferase	0.91	Function unknown	248.47	141.67	351.92	34.86	6.74	14.66
ERBA_02239	TfuA-like protein	0.89	Function unknown	119.34	66.07	163.99	16.46	14.85	24.33
ERBA_00271	Protein of unknown function DUF89	0.83	Function unknown	251.06	92.15	219.19	21.64	13.28	27.77
ERBA_01707	Ribosomal protein S6 modification protein	0.78	Function unknown	341.15	106.73	241.02	133.39	11.23	42.83
ERBA_02240	methanogenesis marker protein 1	0.76	Function unknown	112.58	96.28	214.78	25.16	5.28	9.61
ERBA_01425	hypothetical protein	0.71	Function unknown	298.78	244.94	530.13	48.24	35.67	47.73
ERBA_01557	GTP cyclohydrolase MptA	0.61	Function unknown	602.72	391.15	787.47	45.16	27.05	28.31
ERBA_00544	DNA-directed RNA polymerase subunit F	0.58	Function unknown	974.58	515.5	1019.29	168.94	36.19	98.36
ERBA_00029	hypothetical protein	0.53	Function unknown	726	243.81	468.15	83.98	13.61	58.4
ERBA_01682	**RNA binding protein**	**1.17**	**Function unknown / Coenzyme transport and metabolism**	**136.41**	**108.58**	**321.53**	**13.43**	**15.55**	**9.36**
ERBA_02002	Inosine-5'-monophosphate dehydrogenase	0.52	Function unknown / General function prediction only	158.14	82.13	155.57	7.61	1.73	12.29
ERBA_01990	**Monomethylamine corrinoid protein 1**	**2.28**	**General function prediction only**	**139.97**	**45.22**	**288.89**	**42.81**	**9.66**	**46.25**
ERBA_01443	**TPR-repeat-containing protein**	**1.94**	**General function prediction only**	**72.02**	**32.27**	**159.66**	**15.77**	**17.71**	**23.11**
ERBA_00528	**Met-10+ related protein**	**1.47**	**General function prediction only**	**145.57**	**29.87**	**110.51**	**35.18**	**2.74**	**13.54**
ERBA_00904	**NADPH-dependent FMN reductase protein**	**1.46**	**General function prediction only**	**144.43**	**46.42**	**167.68**	**13.63**	**5.77**	**30.99**
ERBA_01540	**Phosphodiesterase YfcE**	**1.34**	**General function prediction only**	**111.39**	**45.01**	**150.41**	**24.09**	**4.67**	**40.19**
ERBA_00935	**NADPH-dependent FMN reductase protein**	**1.22**	**General function prediction only**	**101.43**	**37.91**	**118.77**	**27.81**	**2.18**	**17.91**
ERBA_00368	**Protein CbbY**	**1.2**	**General function prediction only**	**161.62**	**46.09**	**143.42**	**35.98**	**7.85**	**32.94**
ERBA_00198	**hypothetical protein**	**1.2**	**General function prediction only**	**213.84**	**72.43**	**223.3**	**52.32**	**12.7**	**59.76**
ERBA_01380	**Archaea-specific pyridoxal phosphate-dependent enzyme General function prediction only**	**1.18**	**General function prediction only**	**385.79**	**102.99**	**308.78**	**75.16**	**16.22**	**36.79**
ERBA_00208	**Xylose import ATP-binding protein XylG**	**1.17**	**General function prediction only**	**523.94**	**334.48**	**994.43**	**38.79**	**19.78**	**81.62**
ERBA_00247	**beta-lactamase-like protein**	**1.07**	**General function prediction only**	**779.4**	**316.07**	**879.89**	**186.8**	**3.53**	**32.86**
ERBA_00399	**tRNA 2-agmatinylcytidine synthetase**	**1.07**	**General function prediction only**	**140.32**	**95.77**	**266.16**	**7.9**	**5.39**	**46.29**
ERBA_01550	**ATP binding protein**	**1.04**	**General function prediction only**	**64.97**	**32.25**	**87.93**	**9.62**	**11.35**	**6.18**
ERBA_00807	**2-succinyl-6-hydroxy-2 4-cyclohexadiene-1-carboxylate synthase**	**1.04**	**General function prediction only**	**449.28**	**91.92**	**251.02**	**59.08**	**19.43**	**7.98**
ERBA_00895	small GTP-binding protein	0.98	General function prediction only	167.89	70.88	186.14	10.95	5.93	25.47
ERBA_01922	ABC transporter protein	0.96	General function prediction only	399.03	114.58	296.28	49.73	24.65	36.57
ERBA_01422	putative ribonuclease VapC4	0.87	General function prediction only	276.79	192.6	464.56	48.81	22.7	40.23
ERBA_01693	Octanoyltransferase	0.82	General function prediction only	218.1	105.6	248.73	38.27	14.1	48.72
ERBA_00353	phosphoesterase	0.78	General function prediction only	96.29	51.65	118.08	8.99	13.89	19.5
ERBA_01221	Putative ski2-type helicase	0.76	General function prediction only	79.87	42.78	95.97	11.27	4.04	6.53
ERBA_00246	putative phosphatase-phosphohexomutase General function prediction only	0.76	General function prediction only	266.01	151.95	339.62	51.35	11.54	4.89
ERBA_00158	Memo-like protein	0.75	General function prediction only	1552.73	608.82	1348.99	298.25	32.83	48.55
ERBA_00538	metallophosphoesterase	0.72	General function prediction only	254.71	114.85	252.63	28.84	7.3	48.29
ERBA_01857	Mannosyl-3-phosphoglycerate phosphatase	0.71	General function prediction only	171.14	116.16	253.08	14.65	33.95	36.56
ERBA_00325	metal-dependent phosphohydrolase	0.68	General function prediction only	93.84	106.48	225.31	5.81	18.07	11.47
ERBA_00162	Ribonuclease J	0.6	General function prediction only	482.41	226.26	455.63	137.57	20.82	44.49
ERBA_00411	putative manganese-dependent	0.57	General function prediction only	346.23	141.34	279.67	66.07	9.57	22.91

	inorganic pyrophosphatase							
ERBA_02341	putative cobalt-precorrin-3B C-methyltransferase	0.62	General function prediction only / Amino acid transport and metabolism	385.46	301.93	608.77	59.8	60.14	51.36
ERBA_01592	**Dimethylamine corrinoid protein 1**	**1.44**	**General function prediction only / Coenzyme transport and metabolism**	**68.16**	**14.56**	**51.83**	**24.83**	**3.62**	**14.46**
ERBA_02147	**Ferrous iron transport protein B**	**4.18**	**Inorganic ion transport and metabolism**	**4238.96**	**163.13**	**3907.05**	**252.13**	**9.95**	**279.99**
ERBA_00581	**Zinc transporter ZupT**	**1.24**	**Inorganic ion transport and metabolism**	**123.59**	**17.89**	**56.36**	**29.32**	**7.29**	**17.68**
ERBA_01511	**Cobalamin import system permease protein BtuC**	**1.22**	**Inorganic ion transport and metabolism**	**87.36**	**25.08**	**78.5**	**46.5**	**2.37**	**15.98**
ERBA_00995	**periplasmic copper-binding protein**	**1.02**	**Inorganic ion transport and metabolism**	**33.68**	**20.4**	**55.22**	**7.59**	**4.53**	**3.62**
ERBA_01796	Nitrate transport ATP-binding protein NrtC	0.73	Inorganic ion transport and metabolism	153.4	107.31	233.06	28.8	11.03	37.29
ERBA_00467	**Signal recognition particle 54 kDa protein**	**1.09**	**Intracellular trafficking, secretion, and vesicular transport**	**570.72**	**225.97**	**637.11**	**133.24**	**25.67**	**57.93**
ERBA_01822	Protein-export membrane protein SecF	0.72	Intracellular trafficking, secretion, and vesicular transport	483.9	217.09	473.48	83.69	9.24	30.78
ERBA_01823	Protein-export membrane protein SecD	0.61	Intracellular trafficking, secretion, and vesicular transport	590.41	326.62	656.9	84.36	21.38	32.15
ERBA_00159	Mevalonate kinase	0.92	Lipid transport and metabolism	833.12	261.94	656.37	195.89	21.35	67.87
ERBA_01783	Hydantoin utilization protein A	0.96	Multiple classes	307.62	227.29	584.05	40.16	5.09	27.33
ERBA_01591	Hydantoin utilization protein A	0.84	Multiple classes	47.05	24.19	57.38	19.15	4.62	6.27
ERBA_01861	Hydantoin utilization protein A	0.79	Multiple classes	72.33	42.13	97.04	19.35	4.88	7.82
ERBA_01234	3-isopropylmalate dehydrogenase	0.65	Multiple classes	271.27	179.22	368.94	14.56	15.53	35.09
ERBA_01595	**putative deoxyuridine 5'-triphosphate nucleotidohydrolase**	**1.56**	**Nucleotide transport and metabolism**	**91.22**	**43.66**	**169.49**	**24.54**	**6.78**	**26.39**
ERBA_00565	**Non-canonical purine NTP pyrophosphatase**	**1.56**	**Nucleotide transport and metabolism**	**118.7**	**45.22**	**179.07**	**13.3**	**18.57**	**28.66**
ERBA_00468	**GMP synthase glutamine-hydrolyzing subunit A**	**1.08**	**Nucleotide transport and metabolism**	**249.75**	**95.41**	**266.53**	**57.15**	**7.53**	**38.46**
ERBA_00329	phosphoribosylformylglycinamidine synthase subunit PurS	0.97	Nucleotide transport and metabolism	182.56	112.97	294.83	69.8	8.41	39.25
ERBA_00707	Adenylosuccinate lyase	0.85	Nucleotide transport and metabolism	205.89	98.1	233.29	41.44	10.97	20.13
ERBA_00812	Phosphoribosylamine--glycine ligase	0.76	Nucleotide transport and metabolism	235.56	146.44	330.13	55.3	17.96	58.7
ERBA_02191	Uridylate kinase	0.67	Nucleotide transport and metabolism	447.02	250.3	526.53	46.67	13.61	22.75
ERBA_00122	CTP synthase	0.66	Nucleotide transport and metabolism	386.65	246.53	514.68	17.52	23.93	60.51
ERBA_01553	GMP synthase glutamine-hydrolyzing subunit B	0.65	Nucleotide transport and metabolism	254.36	146.98	303.18	27.88	32.7	42.87
ERBA_01018	**FKBP-type peptidyl-prolyl cis-trans isomerase 2**	**2.59**	**Posttranslational modification, protein turnover, chaperones**	**423.73**	**73.32**	**583.02**	**226.13**	**13.56**	**49.71**
ERBA_00563	**Posttranslational modification putative tRNA threonylcarbamoyladenosine biosynthesis protein KAE1**	**2.38**	**Posttranslational modification, protein turnover, chaperones**	**161.06**	**26.2**	**181**	**36.72**	**4.18**	**9.29**
ERBA_00383	**FKBP-type peptidyl-prolyl cis-trans isomerase 2**	**1.62**	**Posttranslational modification, protein turnover, chaperones**	**17.79**	**14.81**	**61.88**	**8.25**	**1.83**	**7.51**
ERBA_01694	**Posttranslational modification Proteasome-activating nucleotidase**	**1.52**	**Posttranslational modification, protein turnover, chaperones**	**326.16**	**120.24**	**457.02**	**22.36**	**10.42**	**22.72**
ERBA_00809	**Cytochrome c-type biogenesis protein CcmF**	**1.49**	**Posttranslational modification, protein turnover, chaperones**	**104.95**	**40.69**	**152.44**	**16.19**	**6.68**	**29.22**
ERBA_00404	**Chaperone protein ClpB 2**	**1.4**	**Posttranslational modification, protein turnover, chaperones**	**321.98**	**81.86**	**284.97**	**64.38**	**7.03**	**71.9**
ERBA_00197	**Protein-L-isoaspartate O-methyltransferase 1**	**1.35**	**Posttranslational modification, protein turnover, chaperones**	**205.38**	**51.76**	**175.92**	**65.04**	**4.79**	**28.6**
ERBA_00091	**Chaperone protein DnaK**	**1.33**	**Posttranslational modification, protein turnover, chaperones**	**1604.8**	**542.49**	**1796.12**	**133.81**	**13.82**	**238.82**
ERBA_00090	**Protein GrpE**	**1.21**	**Posttranslational modification, protein turnover, chaperones**	**1154.31**	**582.09**	**1775.22**	**183.59**	**39.54**	**58.47**
ERBA_00747	**ADP-ribosylation-Crystallin J1**	**1.04**	**Posttranslational modification, protein turnover, chaperones**	**449.2**	**116.22**	**318.84**	**140.85**	**22.84**	**33.33**
ERBA_01175	**hypothetical protein**	**1.03**	**Posttranslational modification, protein turnover, chaperones**	**224.49**	**43.07**	**116.3**	**56.51**	**4.1**	**11.39**
ERBA_01590	**Thermosome subunit**	**1.03**	**Posttranslational modification, protein turnover, chaperones**	**62.91**	**42.14**	**113.84**	**9.52**	**8.88**	**12.38**
ERBA_01706	Thermosome subunit	1	Posttranslational modification, protein turnover, chaperones	2258.95	1167.8	3072.39	171.91	89.66	92.42
ERBA_00233	Proteasome-activating nucleotidase 1	0.96	Posttranslational modification, protein turnover, chaperones	485	274.75	707.32	22.35	17.16	93.68
ERBA_00092	Chaperone protein DnaJ	0.95	Posttranslational modification, protein turnover, chaperones	994.17	448.48	1141.35	78.9	11.42	57.32
ERBA_01469	Thermosome subunit beta	0.9	Posttranslational modification, protein turnover, chaperones	868.57	588.45	1445.96	135.07	48.83	54.54
ERBA_00825	Archaeal Lon protease	0.81	Posttranslational modification, protein turnover, chaperones	1495.28	429.35	996.41	656.03	16.59	126.99
ERBA_00733	peroxiredoxin 1	0.8	Posttranslational modification, protein turnover, chaperones	284.97	86.07	199.15	34.43	5.95	16.29
ERBA_01845	tRNA-guanine transglycosylase	0.73	Posttranslational modification, protein turnover, chaperones	174.73	138.48	302.58	24.1	22.89	14.71
ERBA_01780	SufBD protein	0.67	Posttranslational modification, protein turnover, chaperones	460.84	335.69	708.15	44.23	8.69	97.35
ERBA_00248	Proteasome subunit beta	0.66	Posttranslational modification, protein turnover, chaperones	1917.54	1064.81	2228.64	159.98	27.85	112.62
ERBA_01943	AIR synthase related protein	0.63	Posttranslational modification, protein turnover, chaperones	402.6	177.55	363.25	99.43	17.13	29.81
ERBA_01781	ABC transporter protein	0.62	Posttranslational modification, protein turnover, chaperones	536.07	325.14	659.64	37.83	17.19	54.14
ERBA_00621	Prefoldin subunit alpha	0.5	Posttranslational modification, protein turnover, chaperones	1013.35	587.56	1099.62	85.86	35.12	93.32
ERBA_00072	**Single-stranded-DNA-specific exonuclease RecJ**	**1.27**	**Replication, recombination and repair**	**418.8**	**119.38**	**381.46**	**92.88**	**1.72**	**46.75**
ERBA_00517	**Transposase**	**1.25**	**Replication, recombination and repair**	**95.82**	**34.88**	**110.53**	**19.49**	**7.9**	**19.91**
ERBA_01295	**Type 2 DNA topoisomerase 6 subunit B**	**1.22**	**Replication, recombination and repair**	**317.69**	**103.06**	**317.68**	**113.45**	**11.78**	**17.24**
ERBA_01220	**Deoxyribodipyrimidine photo-lyase**	**1.11**	**Replication, recombination and repair**	**60.34**	**21.31**	**61.9**	**27.3**	**5.17**	**12.16**
ERBA_01298	DNA gyrase subunit B	0.9	Replication, recombination and repair	511.73	153.32	378.44	111.33	9.33	67.62
ERBA_00710	UvrABC system protein A	0.88	Replication, recombination and repair	185.39	61.08	149.24	53.76	5.01	16.78
ERBA_02286	**transposase IS4 family protein**	**0.82**	**Replication, recombination and repair**	**175.22**	**44.61**	**104.27**	**56.41**	**7.61**	**7.44**
ERBA_01296	Type 2 DNA topoisomerase 6 subunit A	0.8	Replication, recombination and repair	249.07	126.85	291.78	59.86	15.67	19.17

ERBA_01299	DNA gyrase subunit A	0.78	Replication, recombination and repair	368.9	163.18	370.12	66.66	16.67	50.18
ERBA_01980	DNA polymerase	0.76	Replication, recombination and repair	211.32	155.36	348.01	37.23	13.18	26.74
ERBA_01910	tRNA -dimethyltransferase	0.67	Replication, recombination and repair	114.45	110.62	234.28	10.94	4.45	16.81
ERBA_00638	DNA topoisomerase 1	0.6	Replication, recombination and repair	151.61	114.04	227.53	7.63	6.73	9.93
ERBA_00006	DNA primase	0.56	Replication, recombination and repair	710.2	424.34	828.37	62.23	56.72	20.86
ERBA_02016	**Imidazolonepropionase**	**1.06**	**Secondary metabolites biosynthesis, transport and catabolism**	**735.89**	**123.48**	**340.53**	**264.22**	**21.27**	**36.69**
ERBA_00564	**putative bifunctional tRNA threonylcarbamoyladenosine biosynthesis protein**	**2.05**	**Signal transduction mechanisms**	**59.14**	**21.1**	**119.74**	**10.83**	**11.19**	**4.65**
ERBA_01169	**Aerobic respiration control sensor protein ArcB**	**1.66**	**Signal transduction mechanisms**	**79.32**	**23.92**	**101.13**	**15.27**	**4.04**	**31.23**
ERBA_01386	**C4-dicarboxylate transport sensor protein DctS**	**1.15**	**Signal transduction mechanisms**	**68.84**	**25.16**	**74.08**	**12.46**	**5.28**	**7.38**
ERBA_01145	KaiC	0.93	Signal transduction mechanisms	237.71	77.93	196.46	56.5	6.25	11.68
ERBA_01674	Virulence sensor protein BvgS	0.89	Signal transduction mechanisms	88.51	41.59	102.54	16.34	7.13	0.5
ERBA_01628	Blue-light-activated histidine kinase	0.87	Signal transduction mechanisms	118.3	49.99	121.06	45.71	3.29	10
ERBA_01104	**Transcriptional regulator MntR**	**1.97**	**Transcription**	**2188.97**	**207.1**	**1072.47**	**421.03**	**13.39**	**237.19**
ERBA_00433	**HxlR family transcriptional regulator**	**1.37**	**Transcription**	**1106.65**	**168.72**	**577.78**	**246.52**	**45**	**81.29**
ERBA_01424	DNA-directed RNA polymerase subunit E"	0.95	Transcription	249.52	170.29	431.6	22.88	57.42	58.44
ERBA_01423	Ribonuclease R	0.77	Transcription	649.31	378.91	854.75	71.73	49.11	41.04
ERBA_00688	DNA-directed RNA polymerase subunit B"	0.67	Transcription	1270.77	399.88	841.95	424.71	33.96	53.21
ERBA_02087	Transcription elongation factor Spt5	0.53	Transcription	1138.3	615.83	1177.25	65.1	43.2	53.05
ERBA_02145	**Iron-dependent repressor IdeR**	**3.52**	**Transcription /Inorganic ion transport and metabolism**	**2612.08**	**171.31**	**2594**	**285.18**	**12.44**	**123.28**
ERBA_01402	**putative tRNA-dihydrouridine synthase**	**1.83**	**Transcription /Translation, ribosomal structure and biogenesis**	**447.99**	**134.49**	**634.46**	**1.99**	**7.07**	**36.62**
ERBA_01770	**Ribosomal RNA small subunit methyltransferase F**	**1.94**	**Translation, ribosomal structure and biogenesis**	**49.85**	**7.33**	**37.64**	**31.35**	**0.9**	**2.61**
ERBA_01600	**Glycine--tRNA ligase**	**1.39**	**Translation, ribosomal structure and biogenesis**	**178.78**	**68.59**	**237**	**26**	**2.34**	**31.53**
ERBA_01698	**Ribosome-binding ATPase YchF**	**1.2**	**Translation, ribosomal structure and biogenesis**	**113.74**	**60.57**	**183.53**	**10.13**	**5.47**	**8.68**
ERBA_01421	**Translation initiation factor 2 subunit gamma**	**1.19**	**Translation, ribosomal structure and biogenesis**	**551.43**	**266.51**	**801.53**	**53.82**	**11.55**	**47.18**
ERBA_00070	**Histidine--tRNA ligase**	**1.06**	**Translation, ribosomal structure and biogenesis**	**118.43**	**69.63**	**194.43**	**23.07**	**4.79**	**44.49**
ERBA_00187	**Lysine--tRNA ligase 1**	**1.01**	**Translation, ribosomal structure and biogenesis**	**252.87**	**111.28**	**295.63**	**23.25**	**8.81**	**20.44**
ERBA_00739	CCA-adding enzyme	0.98	Translation, ribosomal structure and biogenesis	78.69	40.84	106.3	14.8	3.43	3.09
ERBA_00157	30S ribosomal protein S2	0.96	Translation, ribosomal structure and biogenesis	1537.2	706.06	1815	190.21	94.71	86.28
ERBA_01621	tRNA-guanine transglycosylase	0.93	Translation, ribosomal structure and biogenesis	1278.28	183.29	463.55	572.7	5.26	46.31
ERBA_01607	O-phosphoserine--tRNA ligase	0.86	Translation, ribosomal structure and biogenesis	304.27	148.42	356.35	73.72	11.52	16.27
ERBA_00300	50S ribosomal protein L10e	0.83	Translation, ribosomal structure and biogenesis	1769.49	623.09	1468.59	194.49	42.52	95.96
ERBA_00545	hypothetical protein	0.82	Translation, ribosomal structure and biogenesis	835.27	372.6	871	45.97	42.79	62.19
ERBA_00876	Cysteine--tRNA ligase	0.78	Translation, ribosomal structure and biogenesis	83.26	87.62	199.82	5.69	5.5	36.75
ERBA_00027	30S ribosomal protein S3Ae	0.75	Translation, ribosomal structure and biogenesis	2509.79	1518.33	3364.11	649.13	96.27	193.87
ERBA_01242	Alanine--tRNA ligase	0.72	Translation, ribosomal structure and biogenesis	297.45	240.17	522.34	25.11	16	40.84
ERBA_00071	30S ribosomal protein S15-S13e	0.71	Translation, ribosomal structure and biogenesis	909.17	587.84	1270.53	115.61	70.95	60.16
ERBA_00496	Aspartate--tRNA ligase	0.67	Translation, ribosomal structure and biogenesis	272.58	216.91	458.74	7.37	28.14	34.45
ERBA_00755	30S ribosomal protein S4	0.66	Translation, ribosomal structure and biogenesis	1472.8	899.49	1871.85	158.25	83.07	133.36
ERBA_00141	Valine--tRNA ligase	0.66	Translation, ribosomal structure and biogenesis	493.79	238.95	497.31	30.11	16.79	31.97
ERBA_00756	30S ribosomal protein S13	0.64	Translation, ribosomal structure and biogenesis	1373.6	802.82	1657.91	205.67	83.25	150.14
ERBA_00754	30S ribosomal protein S11	0.63	Translation, ribosomal structure and biogenesis	1537.04	957.42	1953.41	115.55	48.9	90.61
ERBA_00142	Pyrrolysine--tRNA ligase	0.62	Translation, ribosomal structure and biogenesis	412.04	198.99	405.88	80.83	15.31	22.49
ERBA_01924	mRNA 3-end processing factor	0.61	Translation, ribosomal structure and biogenesis	908	429.48	867.47	93.27	12.79	45.96
ERBA_01426	30S ribosomal protein S24e	0.6	Translation, ribosomal structure and biogenesis	651.44	516.46	1040.31	139.01	38.38	91.39
ERBA_00542	tRNA pseudouridine synthase Pus10	0.6	Translation, ribosomal structure and biogenesis	230.06	142.65	284.15	23.45	14.1	40.17
ERBA_00546	putative ribosomal RNA small subunit methyltransferase A	0.58	Translation, ribosomal structure and biogenesis	469.61	266.05	527.93	57.19	21.92	48.38
ERBA_02088	50S ribosomal protein L11	0.56	Translation, ribosomal structure and biogenesis	1468.62	914.22	1783.7	202.83	76.93	138.72
ERBA_00304	Phenylalanine--tRNA ligase alpha subunit	0.53	Translation, ribosomal structure and biogenesis	283.35	142.04	271.78	50.98	6.78	31.28
ERBA_01072	Aspartyl-glutamyl-tRNA amidotransferase subunit B	0.5	Translation, ribosomal structure and biogenesis	252.13	120.23	224.69	27.85	9.86	5.72
ERBA_00149	30S ribosomal protein S9	0.5	Translation, ribosomal structure and biogenesis	767.21	423.35	792.17	49.53	12.87	74.16
ERBA_00732	**Ribosomal RNA large subunit methyltransferase E**	**1.13**	**Translation, ribosomal structure and biogenesis /General function prediction only**	**947.46**	**290.07**	**842.08**	**224.85**	**4.22**	**72.5**
ERBA_02063	Methionine--tRNA ligase	0.6	Translation, ribosomal structure and biogenesis /General function prediction only	566.8	351.16	702.17	60.95	12.4	10.61
ERBA_00301	Translation initiation factor 2 subunit beta	0.55	Translation, ribosomal structure and biogenesis /General function prediction only	1419.72	535.72	1038.98	72.22	41.54	28.44
ERBA_02146	**hypothetical protein**	**3.95**		**1084.46**	**51.84**	**1083.06**	**310.13**	**12.63**	**31.47**
ERBA_01017	**nucleic acid binding OB-fold tRNA-helicase-type**	**2.64**		**105.82**	**24.27**	**205.98**	**52.91**	**8.99**	**51.18**
ERBA_02092	**transposase OrfB family protein**	**2.45**		**298.76**	**16.49**	**124.6**	**166.23**	**3.74**	**35.87**

ERBA_01625	**monomethylamine permease**	**2**	**96.03**	**18.99**	**101.33**	**31.86**	**5.11**	**47.42**
ERBA_01584	**hypothetical protein**	**1.98**	**98.58**	**19.78**	**106.72**	**33.49**	**8.14**	**33.77**
ERBA_02051	**hypothetical protein**	**1.97**	**938.46**	**209.29**	**1088.35**	**136.87**	**31.44**	**163.66**
ERBA_01923	**hypothetical protein**	**1.94**	**563.61**	**106.45**	**545.8**	**146.47**	**15.91**	**28.16**
ERBA_00155	**tRNA-other**	**1.64**	**224.53**	**110.51**	**464.1**	**37.46**	**45.73**	**30.37**
ERBA_02309	**hypothetical protein**	**1.62**	**447.32**	**62.04**	**254.12**	**275.91**	**14.11**	**41.12**
ERBA_01622	**Monomethylamine methyltransferase MtmB1**	**1.61**	**169.53**	**71.52**	**292.99**	**29.77**	**10.7**	**49.96**
ERBA_01504	**putative membrane protein**	**1.59**	**113.52**	**56.3**	**226.6**	**44.8**	**16.88**	**13.62**
ERBA_00941	**transmembrane protein MttP**	**1.5**	**205.42**	**37.53**	**140.29**	**107.65**	**6.43**	**24.64**
ERBA_01457	**hypothetical protein**	**1.44**	**125.43**	**36.93**	**132.41**	**25.81**	**7.11**	**23.83**
ERBA_00402	**trimethylamine permease**	**1.35**	**146.27**	**27.96**	**96.01**	**19.53**	**7.21**	**15.12**
ERBA_01148	**TM helix protein**	**1.34**	**63.89**	**20.5**	**68.96**	**24.73**	**9.93**	**14.54**
ERBA_02308	**hypothetical protein**	**1.25**	**562.79**	**119.1**	**375.19**	**146.02**	**16.85**	**78.43**
ERBA_01626	**hypothetical protein**	**1.22**	**169.05**	**44.87**	**136.22**	**24.89**	**2.77**	**28.28**
ERBA_00579	**Protoglobin domain containing protein**	**1.2**	**116.61**	**21.3**	**67.05**	**32.22**	**8.38**	**28.89**
ERBA_00227	**sensory transduction histidine kinase**	**1.16**	**54.69**	**30.83**	**91.09**	**20.37**	**8.43**	**16.5**
ERBA_01092	**Peptidase C39 like protein**	**1.13**	**36.24**	**18.3**	**52.6**	**3.58**	**0.42**	**12.09**
ERBA_01091	**hypothetical protein**	**1.05**	**41**	**24.64**	**67.39**	**17.21**	**2.53**	**5.48**
ERBA_01309	**hypothetical protein**	**1.02**	**823.21**	**338.31**	**903.72**	**98.05**	**27.73**	**37.6**
ERBA_00586	Calcium-dependent protease	0.96	92.07	29.4	76.17	42.89	5.1	14.39
ERBA_01651	hypothetical protein	0.91	785.11	349.9	871.62	229.7	68.37	31.6
ERBA_01736	hypothetical protein	0.91	220.23	174.86	433.96	33.94	34.46	77.85
ERBA_01244	hypothetical protein	0.88	156.46	46.72	113.47	35.99	2.55	15.71
ERBA_00527	carboxymuconolactone decarboxylase	0.66	423.56	284.79	594.27	77.24	46.87	58.04
ERBA_01392	CoB--CoM heterodisulfide reductase 2 subunit E	0.64	1443.19	681.15	1399.68	141.83	28.03	32.59
ERBA_02125	hypothetical protein	0.57	341.47	235.42	461.8	28.53	24.94	7.43

Table S5. 307 down-regulated genes at 15% salinity relative to 10% salinity.

(Genes with log2 fold change less than -1 are in bold.)

Gene ID	INDIGO annotation	log2 fold change	COG classification	FPKM avg. 5% NaCl	FPKM avg. 10% NaCl	FPKM avg. 15% NaCl	std.dev for 5% NaCl	std.dev for 10% NaCl	std.dev for 15% NaCl
ERBA_00906	**ABC glycine betaine-L-proline transporter substrate-binding subunit**	**-1.9**	**Amino acid transport and metabolism**	**1643.56**	**2477.78**	**880.51**	**747.91**	**68.49**	**79.99**
ERBA_00616	**Glutathione-binding protein GsiB**	**-1.36**	**Amino acid transport and metabolism**	**85.83**	**126.16**	**65.34**	**9.42**	**14.27**	**9.86**
ERBA_00471	**3-dehydroquinate dehydratase**	**-1.18**	**Amino acid transport and metabolism**	**198.89**	**188.85**	**109.59**	**23.14**	**27.14**	**22.68**
ERBA_00472	**Shikimate dehydrogenase**	**-1.11**	**Amino acid transport and metabolism**	**200.14**	**199.84**	**121.41**	**20.68**	**25.87**	**11.91**
ERBA_01085	**Glutamate synthase large subunit-like protein**	**-1.06**	**Amino acid transport and metabolism**	**250.74**	**300.68**	**190.98**	**30.16**	**13.61**	**7.14**
ERBA_00440	P-protein	-0.87	Amino acid transport and metabolism	100.1	98.17	70.58	6.16	7.34	13.22
ERBA_00982	Putative threonine-phosphate decarboxylase	-0.83	Amino acid transport and metabolism	299.11	337.06	246.91	92.52	5.55	57.9
ERBA_01659	Molybdate-tungstate import ATP-binding protein WtpC	-0.67	Amino acid transport and metabolism	216.78	794.15	661.44	49.77	60.09	52.09
ERBA_00473	T-protein	-0.54	Amino acid transport and metabolism / Translation, ribosomal structure and biogenesis	175.21	156.12	142.54	15.24	27.19	16.52
ERBA_00272	23-bisphosphoglycerate-independent phosphoglycerate mutase	-0.95	Carbohydrate transport and metabolism	85.46	129.54	89.45	10.92	7.52	15.52
ERBA_01151	Phosphoenolpyruvate synthase	-0.56	Carbohydrate transport and metabolism	118.22	87.13	77.65	8.6	3.35	5.7
ERBA_02218	**chromosome partitioning ATPase protein**	**-4.1**	**Cell cycle control, cell division, chromosome partitioning**	**211.12**	**1218.58**	**93.78**	**24.73**	**21.47**	**5.54**
ERBA_01566	Cell division protein FtsZ 2	-0.8	Cell cycle control, cell division, chromosome partitioning	1162.58	1343.08	1022.07	298.95	41.66	111.65
ERBA_02104	Chromosome partition protein Smc	-0.77	Cell cycle control, cell division, chromosome partitioning	170.23	231.92	179.69	30.74	16.38	16.49
ERBA_00835	hypothetical protein	-0.73	Cell cycle control, cell division, chromosome partitioning	408.77	130.85	103.8	28.26	1.22	6.15
ERBA_01043	Cell division protein FtsZ 3	-0.71	Cell cycle control, cell division, chromosome partitioning	373.57	442.96	358.53	38.8	21.92	16.11
ERBA_00607	**Flagellin B2**	**-2.48**	**Cell motility**	**1400.6**	**20573.44**	**4895.24**	**1002.33**	**1149.95**	**568.33**
ERBA_01238	**flagellar protein G**	**-1.79**	**Cell motility**	**145.89**	**182.71**	**69.75**	**31.13**	**23.3**	**9.65**
ERBA_01239	**flagellar protein F**	**-1.68**	**Cell motility**	**209.77**	**331.6**	**135.26**	**40.9**	**33.82**	**8.09**
ERBA_01240	**flagella protein**	**-1.55**	**Cell motility**	**349.05**	**480.35**	**217.46**	**69.34**	**17.17**	**29.81**
ERBA_00287	**flagella protein**	**-1.49**	**Cell motility**	**283.89**	**601.55**	**281.69**	**64.89**	**61.16**	**23.45**
ERBA_00596	**N-acetyltransferase GCN5**	**-1.36**	**Cell wall/membrane/envelope biogenesis**	**33.61**	**63.85**	**33.21**	**9.17**	**17**	**5.25**
ERBA_01454	**Immunogenic protein MPT70 analog**	**-1.14**	**Cell wall/membrane/envelope biogenesis**	**188.66**	**67.44**	**40.39**	**30.5**	**18.37**	**5.02**
ERBA_01412	UDP-N-acetylmuramoylalanine--D-glutamate ligase	-0.92	Cell wall/membrane/envelope biogenesis	346.18	262.96	183.22	61.34	28.76	11.6
ERBA_01775	Macrolide export ATP-binding-permease protein MacB	-0.84	Cell wall/membrane/envelope biogenesis	572.45	401.34	294.61	167.26	46.3	18.12
ERBA_01280	Undecaprenyl-phosphate 4-deoxy-4-formamido-L-arabinose transferase	-0.81	Cell wall/membrane/envelope biogenesis	173.26	230.67	173.47	10.79	43.38	23.88
ERBA_01776	S-layer domain protein	-0.77	Cell wall/membrane/envelope biogenesis	813.61	593.56	460.52	95.44	19.9	35.72
ERBA_01656	hypothetical protein	-0.69	Cell wall/membrane/envelope biogenesis	431.77	511.86	418.33	87.71	52.94	41.55
ERBA_01657	hypothetical protein	-0.65	Cell wall/membrane/envelope biogenesis	529.45	542.59	454.56	81.27	23.82	50.34
ERBA_00993	High-affinity choline transport protein	-0.64	Cell wall/membrane/envelope biogenesis	214.45	166.59	140.83	41.7	14.97	17
ERBA_01272	Lipoprotein-releasing system transmembrane protein LolC	-0.57	Cell wall/membrane/envelope biogenesis	240.59	211.07	187.84	8.29	23.52	4.7
ERBA_01383	Poly-beta-16-N-acetyl-D-glucosamine synthase	-0.56	Cell wall/membrane/envelope biogenesis	327.18	296.71	264.49	4.21	43	40.32
ERBA_01740	Archaeal histone A	-0.93	Chromatin structure and dynamics	248.91	505.9	354.78	81.5	51.54	76.73
ERBA_00843	**Methyl-coenzyme M reductase I operon protein**	**-2.35**	**Coenzyme transport and metabolism**	**1283.62**	**4648.87**	**1202.4**	**288.7**	**432.34**	**289.36**

ERBA_01734	**C Trimethylamine methyltransferase MttB**	**-2.26**	**Coenzyme transport and metabolism**	**4636.38**	**18026.02**	**4963.65**	**1519.12**	**1922.67**	**537.4**
ERBA_00842	**Methyl-coenzyme M reductase subunit gamma**	**-2.23**	**Coenzyme transport and metabolism**	**1624.57**	**5171.23**	**1452.08**	**395.84**	**457.92**	**270.18**
ERBA_00844	**Methyl-coenzyme M reductase II operon protein D**	**-1.87**	**Coenzyme transport and metabolism**	**1797.63**	**5083.93**	**1837.19**	**469.48**	**479.22**	**243.11**
ERBA_00845	**Methyl-coenzyme M reductase II subunit beta**	**-1.66**	**Coenzyme transport and metabolism**	**2256.57**	**5665.47**	**2374.78**	**392.18**	**430**	**336.08**
ERBA_00841	**Methyl-coenzyme M reductase II subunit alpha**	**-1.49**	**Coenzyme transport and metabolism**	**2927.67**	**6754.37**	**3183.01**	**580.41**	**635.4**	**304.46**
ERBA_00143	**Uroporphyrinogen-III decarboxylase Coenzyme metabolism**	**-1.41**	**Coenzyme transport and metabolism**	**3866.56**	**9048.37**	**4478.16**	**1906.06**	**745.08**	**877**
ERBA_00871	**Adenosylhomocysteinase**	**-1.41**	**Coenzyme transport and metabolism**	**4026.17**	**6718.35**	**3340.25**	**734.97**	**649.4**	**317.25**
ERBA_00281	**von Willebrand factor type A**	**-1.31**	**Coenzyme transport and metabolism**	**60.69**	**89.58**	**47.78**	**15.4**	**4.28**	**2.13**
ERBA_00615	**Ubiquinone-menaquinone biosynthesis methyltransferase ubiE**	**-1.27**	**Coenzyme transport and metabolism**	**61.57**	**89.96**	**48.62**	**18.22**	**28.03**	**1.91**
ERBA_00872	**S-adenosylmethionine synthase**	**-1.24**	**Coenzyme transport and metabolism**	**3541.53**	**5601**	**3131.66**	**1100.58**	**580**	**362.01**
ERBA_01669	**Trimethylamine methyltransferase MttB**	**-1.09**	**Coenzyme transport and metabolism**	**2263.83**	**1332.41**	**826.15**	**370.74**	**89.57**	**92.9**
ERBA_01569	**D-alanine--poly ligase subunit 1**	**-1.09**	**Coenzyme transport and metabolism**	**528.11**	**333.58**	**206**	**193.29**	**20.58**	**39.39**
ERBA_00729	**Coenzyme A biosynthesis bifunctional protein CoaBC**	**-1.06**	**Coenzyme transport and metabolism**	**86.02**	**108.9**	**69.68**	**13.9**	**16.34**	**9.7**
ERBA_01893	putative cobalt-precorrin-2 C-methyltransferase	-0.74	Coenzyme transport and metabolism	624.14	417.75	330.06	64.6	37.22	26.3
ERBA_00870	glycine N-methyltransferase	-0.54	Coenzyme transport and metabolism	1002.16	878.27	803.52	141.89	46.1	112.37
ERBA_01894	putative cobalt-precorrin-6Y C-methyltransferasedecarboxylating	-0.53	Coenzyme transport and metabolism	617.8	463.41	422.15	48.46	41.13	7.85
ERBA_01777	Macrolide export ATP-binding-permease protein MacB	-0.88	Defense mechanisms	780.28	436.44	314.47	130.25	26.18	37.22
ERBA_01062	restriction endonuclease	-0.82	Defense mechanisms	86.96	66.97	50.27	17.03	7.45	7.21
ERBA_02222	**CoB--CoM heterodisulfide reductase 1 iron-sulfursubunit A**	**-4.79**	**Energy production and conversion**	**108.59**	**495.57**	**24.2**	**59.76**	**26.63**	**13.09**
ERBA_02223	**Nitrogenase iron protein**	**-4.13**	**Energy production and conversion**	**71.98**	**309.78**	**24.05**	**0.9**	**23.26**	**8.62**
ERBA_00865	**GTPase obg**	**-2.17**	**Energy production and conversion**	**754.43**	**720.47**	**210.58**	**86.9**	**79.49**	**14.76**
ERBA_01265	**Type A flavoprotein fprA**	**-1.69**	**Energy production and conversion**	**313.31**	**194.02**	**78.48**	**46.09**	**20.64**	**14.74**
ERBA_01038	**Putative rubrerythrin**	**-1.21**	**Energy production and conversion**	**282.53**	**464.31**	**267.06**	**125.36**	**37.76**	**46.48**
ERBA_01123	**Aspartate ammonia-lyase**	**-1.17**	**Energy production and conversion**	**69.98**	**33.94**	**19.74**	**27.36**	**9.2**	**2.51**
ERBA_01585	**Phycocyanobilin lyase subunit alpha**	**-1.09**	**Energy production and conversion**	**364.51**	**126.32**	**78.68**	**117.22**	**15.01**	**8.95**
ERBA_00116	**NADH-quinone oxidoreductase subunit N**	**-1.08**	**Energy production and conversion**	**489.79**	**1510.38**	**944.81**	**118.71**	**83.52**	**114.32**
ERBA_00115	NADH-quinone oxidoreductase subunit M	-0.91	Energy production and conversion	456.69	1421.11	1001.57	61.34	115.49	122.88
ERBA_01615	Coenzyme F420-dependent N-methenyltetrahydromethanopterin dehydrogenase Energy production	-0.84	Energy production and conversion	1580.25	3493.89	2572.97	547.82	209.25	268.14
ERBA_00054	V-type ATP synthase subunit E	-0.84	Energy production and conversion	1674.83	1680.02	1236.3	299.92	206.12	69.06
ERBA_01634	formylmethanofuran dehydrogenase subunit D	-0.8	Energy production and conversion	962.19	929.47	703.35	247.95	69.21	51.7
ERBA_00113	NADH-quinone oxidoreductase subunit K	-0.8	Energy production and conversion	391.19	999.73	755.07	7.51	74.05	30.95
ERBA_01679	Pyruvate synthase subunit PorD	-0.79	Energy production and conversion	242.68	480.4	368.93	37.73	65.36	141.24
ERBA_01637	Tungsten-containing formylmethanofuran dehydrogenase 2 subunit C	-0.75	Energy production and conversion	1569	1927.62	1517.4	374.57	87.35	120.64
ERBA_01748	Electron transport complex protein RnfG	-0.74	Energy production and conversion	282.76	671.4	530.76	46.14	39.36	3.63
ERBA_02115	510-methylenetetrahydromethanopterin reductase	-0.71	Energy production and conversion	2814.91	3292.75	2655.53	573.71	152.35	253.34
ERBA_01747	Electron transport complex protein RnfE	-0.7	Energy production and conversion	197.29	462.06	377.22	7.33	42.25	48.99
ERBA_00056	V-type ATP synthase subunit F	-0.67	Energy production and conversion	1567.81	1130.36	932.19	461.86	120.77	193.89
ERBA_01677	Pyruvate synthase subunit PorB	-0.67	Energy production and conversion	623.5	1029.37	859.57	181.41	78.25	146.55
ERBA_01746	Electron transport complex protein RnfA	-0.65	Energy production and conversion	175.18	425.64	358.09	8.36	28.04	20.59
ERBA_01411	Nitrogenase molybdenum-iron protein alpha chain	-0.64	Energy production and conversion	377.11	285.37	241.04	73.63	24.75	8
ERBA_00112	NADH-quinone oxidoreductase chain 6	-0.6	Energy production and conversion	541.79	978.8	850.65	73.88	124.33	111.77
ERBA_00052	V-type ATP synthase subunit I	-0.56	Energy production and	2995.73	2424.36	2167.27	372.81	49.23	64.51

			conversion						
ERBA_02220	**dinitrogenase iron-molybdenum cofactor biosynthesi**	**-6.11**	**Function unknown**	**44.69**	**426.2**	**7.68**	**17.71**	**72.71**	**6.65**
ERBA_02221	**dinitrogenase iron-molybdenum cofactor biosynthesi**	**-5.28**	**Function unknown**	**54.76**	**380.51**	**11.52**	**34.2**	**17.52**	**7.37**
ERBA_02217	**putative ArCR protein**	**-3**	**Function unknown**	**245.5**	**1186**	**195.93**	**39.86**	**41.06**	**5.43**
ERBA_00277	**hypothetical protein**	**-1.52**	**Function unknown**	**870.37**	**2993.22**	**1379.36**	**560.19**	**201.3**	**111.04**
ERBA_00278	**flagellin-like protein**	**-1.23**	**Function unknown**	**165.05**	**178.66**	**101.57**	**41.26**	**45.32**	**16.65**
ERBA_02236	**GYD family protein**	**-1.23**	**Function unknown**	**303.4**	**183.03**	**102.38**	**55.18**	**26.88**	**18.23**
ERBA_02143	**putative ACR protein**	**-1.18**	**Function unknown**	**94.8**	**166.17**	**97.36**	**24.55**	**46.65**	**13.7**
ERBA_00957	**hypothetical protein**	**-1.12**	**Function unknown**	**574.75**	**1166.77**	**710.98**	**207.84**	**25.9**	**45.64**
ERBA_01160	**SCP-like extracellular**	**-1.1**	**Function unknown**	**253.45**	**793.04**	**489.7**	**73.51**	**46.47**	**49.15**
ERBA_01492	**Mitochondrial biogenesis AIM24 protein**	**-1.08**	**Function unknown**	**140.06**	**293.25**	**185.4**	**10.15**	**21.86**	**29.82**
ERBA_01453	**hypothetical protein**	**-1.07**	**Function unknown**	**69.24**	**40.38**	**25.37**	**12.64**	**8.74**	**1.41**
ERBA_01264	**hypothetical protein**	**-1.03**	**Function unknown**	**272.24**	**307.42**	**197.62**	**42.87**	**34.15**	**48.12**
ERBA_01413	**Sirohydrochlorin cobaltochelatase**	**-1.01**	**Function unknown**	**379.6**	**624.24**	**410.02**	**102.16**	**91.54**	**41.95**
ERBA_02103	Segregation	-0.99	Function unknown	160.76	229.29	150.6	36.78	29.91	22.66
ERBA_00245	methyltransferase	-0.88	Function unknown	548.55	715.55	509.65	122.91	97.18	78.06
ERBA_02153	PRC-barrel protein	-0.86	Function unknown	516.87	537.35	382.81	90.77	50.25	124.13
ERBA_01465	hypothetical protein	-0.7	Function unknown	323.52	415.04	335.63	80.64	66.76	28.73
ERBA_01338	Serine-threonine-protein kinase pkn1	-0.7	Function unknown	134.09	104.39	84.44	15.26	5.48	1.78
ERBA_01568	methanogenesis marker protein 8	-0.66	Function unknown	492.02	360.43	301.21	128.65	14.95	4.85
ERBA_00920	transmembrane protein	-0.59	Function unknown	227.23	238.19	209.32	27.76	8.72	29.59
ERBA_01820	GTP	-0.57	Function unknown	523.47	263.23	234.49	43.89	14.51	15.74
ERBA_01976	hypothetical protein	-0.51	Function unknown	408.67	388.28	359.13	28.27	7.83	5.53
ERBA_02219	**HTH DNA-binding protein**	**-4.64**	**General function prediction only**	**165.82**	**1024.3**	**54.8**	**53.78**	**42.19**	**14.69**
ERBA_02224	**hypothetical protein**	**-3.71**	**General function prediction only**	**75.78**	**264.53**	**26.77**	**23.73**	**59.56**	**5.65**
ERBA_02225	**Hydroxyacylglutathione hydrolase**	**-2.58**	**General function prediction only**	**119.27**	**216.22**	**47.9**	**44.66**	**37.6**	**14.26**
ERBA_00864	**Protein of unknown function DUF134**	**-2.41**	**General function prediction only**	**1558.34**	**1791.8**	**443.89**	**157.41**	**96.67**	**16.97**
ERBA_01732	**Trimethylamine corrinoid protein 2**	**-2.32**	**General function prediction only**	**4363.07**	**16109.65**	**4265.62**	**1369.17**	**1785.56**	**365.49**
ERBA_02150	**Translation initiation factor 2 subunit beta**	**-1.5**	**General function prediction only**	**1708.7**	**2977.97**	**1383.2**	**1064.1**	**466.57**	**58.63**
ERBA_02149	**Translation initiation factor 2 subunit beta**	**-1.44**	**General function prediction only**	**2700.53**	**4507.03**	**2197.16**	**1572.62**	**609.37**	**336.27**
ERBA_01728	**Dimethylamine corrinoid protein 2**	**-1.33**	**General function prediction only**	**1297.81**	**3805.85**	**1993.76**	**303.77**	**373.43**	**100.46**
ERBA_02061	**Rubredoxin-like zinc ribbon domain protein**	**-1.31**	**General function prediction only**	**148.24**	**268.39**	**142.95**	**38.95**	**28.42**	**20.38**
ERBA_02154	**phosphoglycolate phosphatase**	**-1.27**	**General function prediction only**	**620.07**	**907.18**	**498.28**	**37.39**	**110.57**	**8.75**
ERBA_01357	**protein of unknown function DUF20 protein**	**-1.22**	**General function prediction only**	**187.33**	**281.02**	**159.56**	**41.65**	**32.16**	**12.73**
ERBA_00007	**hypothetical protein**	**-1.15**	**General function prediction only**	**400.45**	**553.76**	**332.82**	**51.64**	**46.38**	**79.79**
ERBA_01640	**S-layer-related protein**	**-1.07**	**General function prediction only**	**10198.99**	**20069.12**	**12638.49**	**1735.89**	**1715.91**	**1391.46**
ERBA_00530	**ArsR family transcriptional regulator**	**-1.04**	**General function prediction only**	**624.75**	**466.63**	**298.87**	**90.75**	**8.84**	**37.12**
ERBA_01133	Dimethylamine corrinoid protein 2	-0.99	General function prediction only	1360.95	1144.76	765.04	148.97	143.41	126.23
ERBA_00008	Putative ribosomal-protein-alanine acetyltransferase	-0.95	General function prediction only	309.38	237.4	163.76	49.51	10.93	22.18
ERBA_00060	Translation initiation factor 2 subunit beta	-0.9	General function prediction only	1879.26	1862.88	1320.61	648.84	99.22	86.84
ERBA_01663	GTPase Der	-0.85	General function prediction only	982.38	861.55	630.5	81.54	32.99	54.88
ERBA_01246	iron-sulfur flavoprotein	-0.8	General function prediction only	161.63	184.85	139.05	5.92	21.27	18.07
ERBA_00512	putative GTP-binding protein EngB	-0.78	General function prediction only	100.51	152.91	119.06	11.9	22.46	14.76
ERBA_00712	molybdenum cofactor biosynthesis protein	-0.66	General function prediction only	300.7	228.42	192.42	47.36	7.33	36.99
ERBA_01467	protein of unknown function DUF87 protein	-0.58	General function prediction only	342.03	371.61	327.86	24.73	31.7	25.25
ERBA_01844	acyltransferase	-0.57	General function prediction only	359.8	325.7	288.03	54.43	24.42	18.24
ERBA_00033	FMN-dependent NADH-azoreductase	-0.57	General function prediction only	1048.27	797.96	707.14	178.18	40.3	92.79
ERBA_01658	Molybdenum transport protein ModE	-0.53	General function prediction only	387.57	701.8	639.93	67.92	129.03	28.49
ERBA_01124	**putative cation-transporting ATPase F**	**-2.02**	**Inorganic ion transport and metabolism**	**98.34**	**138.92**	**45.1**	**0.53**	**11.83**	**2.62**
ERBA_02232	**Manganese-binding lipoprotein MntA**	**-1.77**	**Inorganic ion transport and metabolism**	**132.72**	**153.76**	**59.92**	**32.95**	**40.09**	**6.55**
ERBA_00860	**Putative antiporter subunit mnhC2**	**-1.7**	**Inorganic ion transport and metabolism**	**28.25**	**158.05**	**64.37**	**10.73**	**5.63**	**11.08**
ERBA_02214	**nickel transport complex NikM subunit transmembrane**	**-1.52**	**Inorganic ion transport and metabolism**	**372.15**	**386.73**	**177.1**	**76.01**	**29.73**	**54.34**
ERBA_00858	**Na antiporter subunit E1**	**-1.3**	**Inorganic ion transport and metabolism**	**72.82**	**178.11**	**94.17**	**8.18**	**19.91**	**14.79**
ERBA_00861	**Putative antiporter subunit mnhB2**	**-1.11**	**Inorganic ion transport and metabolism**	**48.05**	**181.47**	**111.31**	**12.6**	**37.2**	**19.72**
ERBA_01391	**Trk system potassium**	**-1.09**	**Inorganic ion transport**	**119.64**	**95.65**	**59.54**	**15.62**	**10.32**	**15.02**

	uptake protein TrkA		**and metabolism**						
ERBA_01662	**Phosphate-specific transport system accessory protein PhoU**	**-1.03**	**Inorganic ion transport and metabolism**	**281.29**	**189.09**	**121**	**64.76**	**11.97**	**21.51**
ERBA_02233	Zinc uptake system ATP-binding protein ZurA	-0.91	Inorganic ion transport and metabolism	150.11	114.54	79.99	65.55	8.79	4.54
ERBA_02234	Manganese transport system membrane protein MntC	-0.85	Inorganic ion transport and metabolism	68.69	59.65	44.03	13.07	5.43	5.3
ERBA_01661	Molybdate-tungstate-binding protein WtpA	-0.81	Inorganic ion transport and metabolism	1282.54	7705.52	5819.05	416.08	327.71	602.59
ERBA_01979	Phosphate-specific transport system accessory protein PhoU	-0.75	Inorganic ion transport and metabolism	1125.2	1196.2	938.57	169.37	53.06	33.39
ERBA_00019	putative ABC transporter permease protein YqgI	-0.74	Inorganic ion transport and metabolism	298.4	282.5	222.93	37.57	1.69	15.16
ERBA_00503	Cation-efflux pump FieF	-0.73	Inorganic ion transport and metabolism	130.24	121.61	96.89	13.4	12	6.97
ERBA_01410	Nitrogenase iron protein	-0.72	Inorganic ion transport and metabolism	310.74	266.12	212.66	68.03	20.48	17.54
ERBA_00020	Phosphate import ATP-binding protein PstB	-0.54	Inorganic ion transport and metabolism	303.09	365.37	330.34	32.77	28.64	36.42
ERBA_01036	**multiple antibiotic resistance protein**	**-1.65**	**Intracellular trafficking, secretion, and vesicular transport**	**57.53**	**29.5**	**12.07**	**17.5**	**10.03**	**2.65**
ERBA_02046	peptidase S26B signal peptidase	-0.98	Intracellular trafficking, secretion, and vesicular transport	333.19	442.8	298.45	29.82	26	27.78
ERBA_02059	**hydroxymethylglutaryl-CoA synthase**	**-1.47**	**Lipid transport and metabolism**	**183.34**	**310.93**	**148.49**	**53.95**	**8.85**	**18.82**
ERBA_02060	**putative acetyl-CoA acyltransferase**	**-1.44**	**Lipid transport and metabolism**	**189.76**	**322.3**	**157.09**	**67.84**	**28.76**	**18.66**
ERBA_00795	**Chemotaxis protein CheW**	**-3.07**	**Multiple classes**	**134.79**	**1370.69**	**215.67**	**38.13**	**191.11**	**13.78**
ERBA_00794	**Methyl-accepting chemotaxis protein TlpA**	**-2.99**	**Multiple classes**	**227.82**	**1681.66**	**279.58**	**57.87**	**191.63**	**3.14**
ERBA_00798	**Chemotaxis protein CheA**	**-2.84**	**Multiple classes**	**74.83**	**648.57**	**119.51**	**19.22**	**91.04**	**15.3**
ERBA_00797	**Chemotaxis response regulator protein-glutamate methylesterase**	**-2.8**	**Multiple classes**	**108.76**	**874.29**	**165.64**	**31.19**	**125.16**	**12.37**
ERBA_00800	**Chemotaxis protein methyltransferase**	**-2.76**	**Multiple classes**	**36.21**	**345.85**	**66.62**	**4.33**	**41.78**	**15.18**
ERBA_00801	**Flagellar motor switch phosphatase FliY**	**-2.68**	**Multiple classes**	**71.03**	**420.06**	**85.74**	**24.21**	**47.23**	**20.92**
ERBA_00799	**Flagellar motor switch phosphatase FliY**	**-2.65**	**Multiple classes**	**67.91**	**498.98**	**104.82**	**27.25**	**68.96**	**7.82**
ERBA_00332	**type II secretion system protein**	**-2.29**	**Multiple classes**	**40.34**	**149.3**	**39.87**	**8**	**14.31**	**4.95**
ERBA_01236	**type II secretion system protein E**	**-1.95**	**Multiple classes**	**159.93**	**193.04**	**65.75**	**42.46**	**25.06**	**12.89**
ERBA_00333	**type II secretion system protein E**	**-1.91**	**Multiple classes**	**128.11**	**280.63**	**98.27**	**23.17**	**9.97**	**3.45**
ERBA_01237	**Circadian clock protein kinase KaiC**	**-1.81**	**Multiple classes**	**123.61**	**179.14**	**66.25**	**32.43**	**20.92**	**12.77**
ERBA_00859	**Na antiporter subunit D1**	**-1.51**	**Multiple classes**	**70.84**	**210.84**	**97.69**	**18.71**	**20.82**	**6.07**
ERBA_01241	**Phosphate regulon transcriptional regulatory protein PhoB**	**-1.24**	**Multiple classes**	**339.75**	**624.58**	**348.98**	**82.08**	**29.06**	**10.86**
ERBA_01235	**flagellar assembly protein J**	**-1.14**	**Multiple classes**	**93.4**	**110.26**	**65.94**	**28.39**	**10.51**	**5.06**
ERBA_01882	**5-methylthioadenosine-S-adenosylhomocysteine deaminase**	**-1.05**	**Multiple classes**	**192.12**	**117.97**	**74.67**	**36.51**	**5.8**	**15.96**
ERBA_02199	Polar-differentiation response regulator DivK	-0.97	Multiple classes	289.03	442.48	297.68	85.56	14.4	10.51
ERBA_02021	Sodium-proline symporter	-0.94	Multiple classes	135.28	184.34	126.96	29.82	19.11	7.6
ERBA_02049	ORC1-type DNA replication protein 1	-0.78	Multiple classes	477.02	269.49	206.91	13.93	2.95	4.71
ERBA_00347	DNA repair helicase	-0.77	Multiple classes	268.65	200.77	156.92	42.29	16.39	23.76
ERBA_01788	Anthranilate synthase component 1 2	-0.74	Multiple classes	97.32	172.1	136.19	21.17	11.24	18.8
ERBA_00114	NADH-quinone oxidoreductase subunit L	-0.67	Multiple classes	469.79	1265.41	1053.77	88.89	118.71	91.63
ERBA_00862	Na antiporter subunit A1	-0.59	Multiple classes	191.64	343.3	301.79	17.79	21.49	35.84
ERBA_02139	**Phosphoribosylaminoimidazole-succinocarboxamide synthase**	**-1.44**	**Nucleotide transport and metabolism**	**244.86**	**319.59**	**154.85**	**35.27**	**22.13**	**3.24**
ERBA_02062	**adenylate cyclase**	**-1.33**	**Nucleotide transport and metabolism**	**215.05**	**276.2**	**145.82**	**72.76**	**34.83**	**21.62**
ERBA_00839	**Putative thymidylate synthase**	**-1.25**	**Nucleotide transport and metabolism**	**366.27**	**217.93**	**121.49**	**134.13**	**18.52**	**26.61**
ERBA_02282	Deoxyguanosinetriphosphate triphosphohydrolase-like protein	-0.86	Nucleotide transport and metabolism	72.92	159.8	116.66	29.93	21.29	15.92
ERBA_00451	Aspartate carbamoyltransferase	-0.82	Nucleotide transport and metabolism	171.87	285.45	213.12	14.29	29.84	11.22
ERBA_00452	Aspartate carbamoyltransferase regulatory chain	-0.67	Nucleotide transport and metabolism	154.6	192.68	157.95	30.54	9.68	36.29
ERBA_02054	Bifunctional purine biosynthesis protein PurH	-0.6	Nucleotide transport and metabolism	867.01	448.37	391.52	274.39	19.36	6.01
ERBA_01545	Peptide methionine sulfoxide reductase MsrA	-0.97	Posttranslational modification, protein turnover, chaperones	131.56	152.59	104.83	44.18	20.77	36.3
ERBA_01211	Cytochrome c biogenesis protein Posttranslational modification	-0.88	Posttranslational modification, protein turnover, chaperones	178.79	70.77	51.41	22.08	7.28	8.41
ERBA_00274	putative peroxiredoxin	-0.81	Posttranslational	698.75	411.23	310.75	122.64	30.49	38.53

			modification, protein turnover, chaperones						
ERBA_01941	Protein QmcA	-0.8	Posttranslational modification, protein turnover, chaperones	342.9	274.86	208.08	12.78	35.84	7.02
ERBA_01873	Proteasome-activating nucleotidase	-0.68	Posttranslational modification, protein turnover, chaperones	969.45	1045.72	863.17	225.18	47.68	56.49
ERBA_00206	peptidase M10A	-0.67	Posttranslational modification, protein turnover, chaperones	211.72	189.81	157.54	5.06	11.54	22.74
ERBA_01681	ATP-dependent 26S proteasome regulatory subunit Posttranslational modification	-0.54	Posttranslational modification, protein turnover, chaperones	184.22	177.15	160.52	2.22	14.81	2.65
ERBA_01327	**Putative transposase in snaA-snaB intergenic region**	**-1.92**	**Replication, recombination and repair**	**433.45**	**449.69**	**156.98**	**122.21**	**19.77**	**11.26**
ERBA_02155	**DNA repair**	**-1.53**	**Replication, recombination and repair**	**1532.49**	**2721.12**	**1239.42**	**240.25**	**115.56**	**99.01**
ERBA_00723	**CRISPR-associated exonuclease Cas4**	**-1.27**	**Replication, recombination and repair**	**61.68**	**163.44**	**89.86**	**10.41**	**18.82**	**17.18**
ERBA_00724	**CRISPR-associated endoribonuclease Cas6 1**	**-1.24**	**Replication, recombination and repair**	**127.1**	**295.18**	**166.04**	**24.26**	**21.5**	**25.65**
ERBA_00331	**Ribonuclease HII**	**-1.02**	**Replication, recombination and repair**	**83.77**	**110.89**	**71.94**	**12.81**	**32.73**	**6.97**
ERBA_00722	**CRISPR-associated protein Cas1 1**	**-1**	**Replication, recombination and repair**	**88.14**	**154.16**	**101.41**	**24.41**	**14.69**	**12.5**
ERBA_00958	replication factor A	-1	Replication, recombination and repair	857.53	1430.62	944.93	291.24	55.63	27.6
ERBA_02106	DNA double-strand break repair Rad50 ATPase	-0.94	Replication, recombination and repair	141.99	147.44	101.11	57.14	14.77	15.94
ERBA_02107	DNA double-strand break repair protein Mre11	-0.74	Replication, recombination and repair	222.8	181.45	142.49	69.5	16.68	18.82
ERBA_00382	Putative transposase in snaA-snaB intergenic region	-0.74	Replication, recombination and repair	261.62	249.13	197.83	56.26	17.16	9.5
ERBA_01877	Toprim sub domain protein	-0.67	Replication, recombination and repair	439.29	275.35	230.73	59.31	19.58	28.84
ERBA_01601	hypothetical protein	-0.64	Replication, recombination and repair	371.21	389.28	330.22	35.43	26.96	34.01
ERBA_01743	Replication factor C small subunit	-0.59	Replication, recombination and repair	197.48	221.81	196.55	26.58	5.52	31.58
ERBA_02277	DNA helicase II - ATP-dependent DNA helicase PcrA	-0.56	Replication, recombination and repair	78.11	104.17	93.21	3.01	0.35	3.22
ERBA_00015	Tat-linked quality control protein TatD	-0.51	Replication, recombination and repair	288.97	253.99	235.68	17.81	24.7	14.18
ERBA_00796	**Chemotaxis response regulator protein-glutamate methylesterase**	**-3.09**	**Signal transduction mechanisms**	**114.72**	**1049.52**	**162.25**	**11.05**	**143.57**	**5.3**
ERBA_01949	**Polar-differentiation response regulator DivK**	**-2.38**	**Signal transduction mechanisms**	**222.16**	**388.05**	**96.44**	**97.36**	**46.46**	**39.23**
ERBA_00599	**putative circadian clock protein KaiC**	**-1.3**	**Signal transduction mechanisms**	**168.58**	**402.61**	**216.03**	**5.5**	**14.41**	**16.4**
ERBA_01977	**Circadian clock protein kinase KaiC**	**-1.24**	**Signal transduction mechanisms**	**931.25**	**919.61**	**513.03**	**219.08**	**25.05**	**23.86**
ERBA_02035	Circadian clock protein kinase KaiC	-0.83	Signal transduction mechanisms	457.8	849.55	630.97	112.73	49.59	74.98
ERBA_01430	Sensor protein IrlS	-0.81	Signal transduction mechanisms	68.8	73.13	54.59	15.23	3.44	3.86
ERBA_01274	multisensor signal transduction histidine kinase	-0.75	Signal transduction mechanisms	110.38	207.51	162.26	35.69	13.12	17.14
ERBA_00830	Aerobic respiration control sensor protein ArcB	-0.73	Signal transduction mechanisms	434.56	366.04	290.97	9.04	22.79	1.5
ERBA_02058	**XRE family transcriptional regulator**	**-2.1**	**Transcription**	**163.26**	**446.8**	**138.01**	**49.56**	**70.73**	**19.49**
ERBA_02053	**ArsR family transcriptional regulator**	**-1.52**	**Transcription**	**360.71**	**332.4**	**152.54**	**67.42**	**20.17**	**19.11**
ERBA_02231	**Transcriptional regulator MntR**	**-1.18**	**Transcription**	**123.13**	**220.87**	**127.14**	**53.08**	**40.78**	**10.19**
ERBA_01778	**MarR family transcriptional regulator**	**-1.13**	**Transcription**	**897.27**	**617.29**	**373.27**	**183.94**	**20.13**	**33.25**
ERBA_01795	XRE family transcriptional regulator	-0.88	Transcription	79.6	77.86	56.03	14.13	12.45	7.75
ERBA_02007	transcriptional regulator protein	-0.84	Transcription	1128.86	447.93	334.38	165.66	52.12	96.17
ERBA_02102	Segregation	-0.73	Transcription	186.22	171.93	137.09	4.16	28.11	7.27
ERBA_02194	ArsR family transcriptional regulator	-0.68	Transcription	485.08	405.22	336.76	71.92	43.21	75.96
ERBA_02334	transcriptional regulator TrmB	-0.68	Transcription	334.55	322.52	264.48	59.32	23.15	21.66
ERBA_01934	Helix-turn-helix domain	-0.56	Transcription	891.28	681.63	608.4	174.78	34.94	32.58

	containing protein								
ERBA_02281	**tRNA-guanine transglycosylase**	**-2.8**	**Translation, ribosomal structure and biogenesis**	**279.99**	**238.67**	**44.93**	**82.05**	**26.09**	**7.43**
ERBA_00166	**H-ACA RNA-protein complex component Gar1**	**-1.49**	**Translation, ribosomal structure and biogenesis**	**939.96**	**869.36**	**404.56**	**121.43**	**128.71**	**157.57**
ERBA_00782	30S ribosomal protein S3	-0.89	Translation, ribosomal structure and biogenesis	815.36	675.63	479.74	242.39	152.07	156.21
ERBA_00783	50S ribosomal protein L22P	-0.79	Translation, ribosomal structure and biogenesis	982.5	612.76	468.36	259.95	116.3	134.4
ERBA_00678	30S ribosomal protein S10	-0.69	Translation, ribosomal structure and biogenesis	1624.47	1901.5	1553.8	322.37	169.87	103.84
ERBA_00767	50S ribosomal protein L15P	-0.69	Translation, ribosomal structure and biogenesis	547.93	669.47	550.44	85.41	105.66	99.75
ERBA_00785	50S ribosomal protein L2	-0.6	Translation, ribosomal structure and biogenesis	1594.9	745.86	648.76	422.82	65.79	89.46
ERBA_00252	NMD3 family protein	-0.51	Translation, ribosomal structure and biogenesis	442.35	634.57	589.48	33.77	32.47	44.65
ERBA_00706	**hypothetical protein**	**-3**		**43.34**	**64.94**	**8.19**	**23.02**	**11.77**	**7.13**
ERBA_00802	**hypothetical protein**	**-2.3**		**57.12**	**304.96**	**82.01**	**11.1**	**35.87**	**7.84**
ERBA_00585	**hypothetical protein**	**-2.15**		**97.78**	**120.37**	**35.93**	**55.78**	**39.7**	**11.84**
ERBA_01556	**chromosomal protein**	**-2.05**		**938.59**	**3846.93**	**1222.96**	**563.34**	**312.47**	**67.47**
ERBA_00285	**hypothetical protein**	**-2.05**		**47.97**	**155.11**	**50.53**	**14.96**	**27.67**	**16.2**
ERBA_00286	**hypothetical protein**	**-2.05**		**65.79**	**270.87**	**87.92**	**7.81**	**29.02**	**25.07**
ERBA_02082	**hypothetical protein**	**-2**		**46.04**	**35.96**	**12.66**	**10.51**	**9.08**	**12.15**
ERBA_00130	**hypothetical protein**	**-1.89**		**6705.14**	**6076.1**	**2165.6**	**4924.31**	**501.79**	**178.8**
ERBA_00276	**hypothetical protein**	**-1.87**		**510.29**	**2788.5**	**1016.58**	**389.74**	**101.54**	**255.11**
ERBA_01575	**hypothetical protein**	**-1.86**		**290.15**	**287.35**	**106.07**	**56.18**	**21.4**	**40.54**
ERBA_01582	**EF hand protein**	**-1.82**		**1966.31**	**775.33**	**290.78**	**1412.93**	**199.63**	**38.35**
ERBA_00484	**hypothetical protein**	**-1.82**		**981.2**	**410.93**	**153.61**	**720.41**	**87.87**	**28.71**
ERBA_00605	**hypothetical protein**	**-1.79**		**61.47**	**173.99**	**67**	**20.59**	**27.84**	**12.86**
ERBA_02283	**hypothetical protein**	**-1.77**		**111.56**	**426.34**	**167.44**	**42.11**	**44.59**	**36.26**
ERBA_00448	**hypothetical protein**	**-1.73**		**172.06**	**268.51**	**104.88**	**50.56**	**72.78**	**41.26**
ERBA_00597	**hypothetical protein**	**-1.69**		**73.19**	**227.07**	**92.02**	**12.43**	**6.18**	**6.99**
ERBA_02288	**hypothetical protein**	**-1.69**		**156.11**	**221.65**	**88.43**	**29.76**	**16.89**	**15.01**
ERBA_02347	**transposase IS4**	**-1.63**		**41770.91**	**65261.13**	**27812.42**	**9789.21**	**8862.67**	**2390.01**
ERBA_00284	**hypothetical protein**	**-1.59**		**40.96**	**125.05**	**56.22**	**11.76**	**15.21**	**18.31**
ERBA_01730	**hypothetical protein**	**-1.57**		**1607.86**	**4747.66**	**2108.9**	**111.43**	**473.84**	**418.8**
ERBA_00511	**protein of unknown function DUF11 protein**	**-1.56**		**136.86**	**217.8**	**96.62**	**28.26**	**20.25**	**24.38**
ERBA_00600	**hypothetical protein**	**-1.52**		**141.03**	**315.75**	**143.98**	**17.39**	**42.34**	**11.7**
ERBA_01731	**trimethylamine permease**	**-1.49**		**1573.26**	**5105.33**	**2408.13**	**65.24**	**345.85**	**437.51**
ERBA_00598	**hypothetical protein**	**-1.47**		**59.2**	**196.81**	**92.5**	**8.6**	**6.9**	**10.01**
ERBA_01581	**hypothetical protein**	**-1.46**		**2381.24**	**909.48**	**436.73**	**1653.64**	**226.65**	**62.55**
ERBA_00561	**small heat shock protein**	**-1.45**		**1178.21**	**2558.25**	**1230.78**	**308.12**	**155.57**	**49.26**
ERBA_00053	**putative ATPase proteolipid chain**	**-1.45**		**2306.78**	**3798.97**	**1829.98**	**376.48**	**121.76**	**60.71**
ERBA_02025	**hypothetical protein**	**-1.4**		**156.06**	**388.36**	**193.72**	**86.04**	**68.83**	**11.43**
ERBA_00603	**N-acetyltransferase GCN5**	**-1.39**		**67.27**	**90.02**	**44.76**	**42.48**	**9.2**	**13**
ERBA_01729	**hypothetical protein**	**-1.39**		**1013.74**	**2872.94**	**1442.15**	**122.58**	**118.07**	**241.79**
ERBA_01919	**hypothetical protein**	**-1.38**		**56.74**	**151.1**	**75.11**	**24.52**	**20.88**	**17.46**
ERBA_01090	**hypothetical protein**	**-1.35**		**72.1**	**166.72**	**87.17**	**24.51**	**15.41**	**12.89**
ERBA_00601	**hypothetical protein**	**-1.34**		**179.05**	**436.76**	**227.85**	**25.71**	**12.04**	**44.84**
ERBA_01486	**Dimethylamine methyltransferase MtbB1**	**-1.3**		**152.38**	**304.57**	**164.48**	**60.48**	**59.77**	**25.72**
ERBA_01829	**hypothetical protein**	**-1.3**		**643.83**	**935.5**	**505.32**	**302.96**	**46.21**	**69.96**
ERBA_00283	**hypothetical protein**	**-1.28**		**38.73**	**84.06**	**44.69**	**5.74**	**19.37**	**11.06**
ERBA_00606	**hypothetical protein**	**-1.28**		**112.01**	**664.87**	**361.47**	**43.83**	**59.03**	**62.49**
ERBA_01921	**TM2 domain-containing protein**	**-1.27**		**701.53**	**535.28**	**297.56**	**230.48**	**29.99**	**90.11**
ERBA_02312	**hypothetical protein**	**-1.26**		**822.79**	**483.94**	**268.42**	**76.62**	**65.47**	**17.09**
ERBA_00117	**F dehydrogenase subunit FpoO**	**-1.26**		**506.11**	**1508.24**	**832.7**	**125.58**	**135.01**	**149.83**
ERBA_00804	**hypothetical protein**	**-1.24**		**223.74**	**613.38**	**342.85**	**69.67**	**46.85**	**21.2**
ERBA_02066	**hypothetical protein**	**-1.23**		**369.22**	**309.26**	**171.64**	**74.01**	**49.66**	**55.21**
ERBA_00602	**hypothetical protein**	**-1.21**		**205.81**	**425.53**	**242.53**	**45.28**	**10.28**	**8.15**
ERBA_00803	**hypothetical protein**	**-1.19**		**250**	**614.46**	**356.28**	**74.3**	**41.96**	**26.09**
ERBA_01227	**hypothetical protein**	**-1.17**		**162.06**	**508.32**	**297.23**	**30.07**	**108.62**	**58.1**
ERBA_00558	**hypothetical protein**	**-1.14**		**276.69**	**223.92**	**135.07**	**15.57**	**13.4**	**14.14**
ERBA_00282	**hypothetical protein**	**-1.13**		**54.74**	**79.76**	**47.86**	**7.62**	**8.42**	**12.02**
ERBA_02009	**hypothetical protein**	**-1.13**		**389.98**	**308.7**	**185.78**	**123.09**	**19.24**	**23.42**
ERBA_00731	**hypothetical protein**	**-1.12**		**568.65**	**733.66**	**445.53**	**136.82**	**41.15**	**39.08**
ERBA_02343	**Methylated-thiol--corrinoid protein MtsB**	**-1.09**		**1276.53**	**461.31**	**285.74**	**149.17**	**22.45**	**51.84**
ERBA_01207	**methyltransferase FkbM**	**-1.07**		**174.46**	**218.43**	**137.4**	**49.11**	**8.29**	**30.52**
ERBA_01212	**hypothetical protein**	**-1.06**		**177.81**	**117.54**	**76.34**	**7.01**	**19.8**	**26.44**
ERBA_01442	**hypothetical protein**	**-1.01**		**365.51**	**262.25**	**171.57**	**68.75**	**65.23**	**24.17**
ERBA_01359	**hypothetical protein**	**-1**		**78.2**	**63.42**	**41.61**	**19.51**	**21.01**	**7.55**
ERBA_00510	**protein of unknown function DUF11 protein**	**-1**		**232.98**	**261.06**	**171.88**	**43.76**	**10.23**	**10.16**
ERBA_02287	hypothetical protein	-0.99		112.5	97.76	64.28	26.38	20.51	11.98
ERBA_00550	hypothetical protein	-0.99		169.78	288.65	190.98	21.09	31.11	33.7
ERBA_01918	Translocon-associated beta protein	-0.98		141.05	351.86	236.56	9.26	9.12	24
ERBA_00534	cell surface lipoprotein	-0.95		482.8	744.13	507.21	150.87	68.06	54.79
ERBA_02278	PD-D/EXK nuclease superfamily protein	-0.95		115.53	161.28	109.64	18.43	16.15	10.9
ERBA_00960	Dimethylamine methyltransferase MtbB3	-0.93		3300.05	5914.41	4102.77	409.34	381.35	190.92
ERBA_00504	hypothetical protein	-0.93		146.76	271.28	187.79	42.7	13.52	14.88
ERBA_00591	Phospholipase D-nuclease domain containing protein	-0.92		237.85	326.61	228.97	55.34	33.83	56.7
ERBA_02175	hypothetical protein	-0.9		211.92	148.83	105.42	25.25	21.46	2.26
ERBA_00051	H-transporting ATP synthase	-0.89		2518.8	2504.06	1782.85	178.95	284.09	180.61

	subunit H							
ERBA_01134	Monomethylamine methyltransferase MtmB2	-0.89	733.57	1581.22	1129.24	322.28	57.71	69.49
ERBA_02159	Winged helix-turn-helix DNA-binding protein	-0.87	810.91	170.98	124.68	284.7	8.56	17.53
ERBA_00929	CHAD domain containing protein	-0.87	336.59	722.94	523.82	28.79	54.14	47.34
ERBA_01498	hypothetical protein	-0.86	308.54	86.34	63.29	95.83	5.5	15.94
ERBA_00930	hypothetical protein	-0.86	796.65	1372.44	1000.98	86.5	18.33	74.93
ERBA_01360	hypothetical protein	-0.86	124.51	95.78	71.06	13.55	14.74	19.4
ERBA_02055	Putative adenylate cyclase 3	-0.85	846.39	1652.25	1212.22	179.4	84.54	155.73
ERBA_01772	hypothetical protein	-0.83	159.9	117.17	86.57	30.35	14.97	2.79
ERBA_00882	methanol protein	-0.83	1318.86	598.12	444.76	59.45	55.3	51.23
ERBA_01044	conserved repeat domain protein	-0.82	211.81	301.59	224.76	5.16	27.21	19.36
ERBA_01794	S-layer-related duplication domain protein	-0.8	363.43	417.34	316.88	44.17	39.2	26.52
ERBA_00529	ArsR family transcriptional regulator	-0.79	1093.88	1058.64	808.88	333.55	44.7	60.96
ERBA_02289	hypothetical protein	-0.75	150.71	310.04	242.85	27.79	11.43	46.29
ERBA_01950	hypothetical protein	-0.73	279.76	314.36	250.11	79.74	32.25	12.51
ERBA_00441	hypothetical protein	-0.73	104.28	116.94	94	6.38	17.75	11.55
ERBA_01988	hypothetical protein	-0.72	436.14	523.71	415.68	37.54	50.46	59.11
ERBA_01665	hypothetical protein	-0.71	582.35	638.5	514.78	26.1	38.24	9.02
ERBA_00234	hypothetical protein	-0.71	2861.91	1292.7	1042.97	633.27	52.95	81.35
ERBA_01366	hypothetical protein	-0.7	109.35	76.35	62.99	14.84	9.83	15.25
ERBA_01638	DNA polymerase beta subunit	-0.67	570.46	583.27	480.14	76.53	22.61	35.22
ERBA_00956	hypothetical protein	-0.64	312.37	485.21	412.93	30.14	65.87	43.08
ERBA_02034	hypothetical protein	-0.52	278.1	446.8	416.05	10.22	21.14	51.33

Table S6. 433 up-regulated genes at 5% salinity relative to 10% salinity. (Genes with log2 fold change greater than 1 are in bold.)

Gene ID	INDIGO annotation	log2 fold change	COG classification	FPKM avg. 5% NaCl	FPKM avg. 10% NaCl	FPKM avg. 15% NaCl	std.dev for 5% NaCl	std.dev for 10% NaCl	std.dev for 15% NaCl
ERBA_02242	**putative cysteine desulfurase**	**2.95**	**Amino acid transport and metabolism**	**2073.18**	**180.67**	**488.52**	**1009.17**	**12.94**	**77.68**
ERBA_00486	**dienelactone hydrolase**	**2.2**	**Amino acid transport and metabolism**	**409.13**	**59.81**	**261.9**	**207.83**	**12.3**	**43.42**
ERBA_01548	**putative cysteine desulfurase**	**2.09**	**Amino acid transport and metabolism**	**169.39**	**26.71**	**51.72**	**99.23**	**2.24**	**8.8**
ERBA_02050	**L-tyrosine decarboxylase**	**2**	**Amino acid transport and metabolism**	**862.86**	**146.71**	**488.3**	**266.3**	**14.41**	**56.17**
ERBA_00168	**lysine exporter protein LysE-YggA**	**1.93**	**Amino acid transport and metabolism**	**230.98**	**40.47**	**65.91**	**103.54**	**17.64**	**19.97**
ERBA_01049	**S-adenosylmethionine synthase**	**1.77**	**Amino acid transport and metabolism**	**215.33**	**43.3**	**86.12**	**35.16**	**2.17**	**12.13**
ERBA_01555	**Acetylglutamate kinase**	**1.15**	**Amino acid transport and metabolism**	**159.88**	**49.27**	**211.97**	**32.13**	**11.05**	**48.4**
ERBA_01928	**510-methylenetetrahydrofolate reductase**	**1.13**	**Amino acid transport and metabolism**	**416.13**	**128.76**	**172.39**	**177.6**	**10.17**	**28.09**
ERBA_00806	**putative amino acid-proton symporter YbeC**	**1.1**	**Amino acid transport and metabolism**	**123.92**	**39.42**	**68.78**	**39.34**	**3.66**	**11.77**
ERBA_01098	**NADPH-dependent glutamate synthase large subunit**	**1.07**	**Amino acid transport and metabolism**	**136.98**	**44.32**	**117.33**	**41.15**	**10.05**	**33.36**
ERBA_00269	**Putative fructoselysine transporter FrlA**	**1.03**	**Amino acid transport and metabolism**	**143.08**	**48.2**	**71.79**	**20.05**	**6.67**	**9.03**
ERBA_02079	Choline transport system permease protein OpuBB	0.95	Amino acid transport and metabolism	186.61	65	113.64	66.27	23.5	35
ERBA_01381	Cystathionine gamma-synthase	0.95	Amino acid transport and metabolism	219.03	77.45	202.33	51.27	15.9	21.47
ERBA_01624	D-serine-D-alanine-glycine transporter	0.93	Amino acid transport and metabolism	129.65	46.39	145.28	19.21	11.12	22.93
ERBA_01841	Acetylornithine aminotransferase	0.88	Amino acid transport and metabolism	318.21	120.05	261.62	31.23	4.94	38.72
ERBA_01042	putative amino acid-proton symporter YbeC	0.84	Amino acid transport and metabolism	103.41	39.08	37.82	40.99	8.06	2.87
ERBA_00619	Tryptophan synthase alpha chain	0.79	Amino acid transport and metabolism	294.29	115.95	153.5	90.19	0.78	8.19
ERBA_00714	putative 2-isopropylmalate synthase	0.78	Amino acid transport and metabolism	845.85	338.75	612.56	85.4	10.96	51.89
ERBA_01030	3-isopropylmalate dehydratase large subunit 1	0.68	Amino acid transport and metabolism	306.18	130.43	270.31	80.08	16.44	40.72
ERBA_00823	threonine synthase	0.65	Amino acid transport and metabolism	489.61	212.62	396.52	115.52	5.59	15.05
ERBA_01099	CoB--CoM heterodisulfide reductase 1 iron-sulfursubunit A	0.96	Amino acid transport and metabolism /Energy production and conversion	134.25	46.61	131.83	48.28	14.33	23.62
ERBA_01382	Homoserine O-acetyltransferase	0.55	Amino acid transport and metabolism /General function prediction only	171.29	80.48	187	40.8	7.13	7.04
ERBA_00431	**cupin**	**2.22**	**Carbohydrate transport and metabolism**	**405.05**	**59.45**	**446.87**	**64.5**	**10.75**	**58.36**
ERBA_00035	**glycoside hydrolase 15-like protein**	**1.96**	**Carbohydrate transport and metabolism**	**964.5**	**167.69**	**394.65**	**333.62**	**21.79**	**88.08**
ERBA_00974	**transporter**	**1.52**	**Carbohydrate transport and metabolism**	**74.69**	**17.37**	**28.96**	**44.92**	**0.48**	**4.91**
ERBA_02301	**Tetracycline resistance protein class G**	**1.47**	**Carbohydrate transport and metabolism**	**195.72**	**47.75**	**95.92**	**90.32**	**4.8**	**4.54**
ERBA_00867	**Tetracycline resistance protein class E**	**1.4**	**Carbohydrate transport and metabolism**	**260.42**	**66.92**	**96.39**	**106.04**	**12.22**	**14.21**
ERBA_01431	**multidrug-efflux transporter**	**1.39**	**Carbohydrate transport and metabolism**	**91.28**	**23.41**	**50.59**	**51.38**	**0.9**	**5.59**
ERBA_00640	**Methyl-coenzyme M reductase II subunit beta**	**1.22**	**Carbohydrate transport and metabolism**	**109.44**	**32.3**	**59.85**	**12.1**	**10.05**	**10.13**
ERBA_02193	putative endonuclease 4	0.86	Carbohydrate transport and metabolism	166.73	62.23	119.14	57.5	2.23	36.98
ERBA_00937	quinoprotein glucose dehydrogenase	0.86	Carbohydrate transport and metabolism	103.3	38.68	81.61	28.01	5.77	20.68
ERBA_00119	Fructose-16-bisphosphatase FBPase V	0.78	Carbohydrate transport and metabolism	78.8	31.55	149.27	5.11	1.94	28.59
ERBA_01868	putative phosphoglucosamine mutase	0.58	Carbohydrate transport and metabolism	259.49	119.26	160.83	34.53	11.04	21.93
ERBA_00221	23-bisphosphoglycerate-independent phosphoglycerate mutase 1	0.53	Carbohydrate transport and metabolism	271.86	130.07	273.69	23.44	9.31	15.49
ERBA_00295	**Protein CrcB**	**2.72**	**Cell cycle control, cell division, chromosome partitioning**	**174.76**	**17.53**	**43.21**	**82.64**	**2.36**	**34.47**
ERBA_00434	**PP-loop domain containing protein**	**2.48**	**Cell cycle control, cell division, chromosome partitioning**	**725.07**	**89.16**	**311.31**	**127.18**	**3.32**	**25.08**
ERBA_00294	**Protein CrcB**	**2.42**	**Cell cycle control, cell**	**163.8**	**20.04**	**44.36**	**107.45**	**6.96**	**6.57**

ERBA_00835	**hypothetical protein**	**1.1**	**Cell cycle control, cell division, chromosome partitioning**	**408.77**	**130.85**	**103.8**	**28.26**	**1.22**	**6.15**
ERBA_02014	**Small-conductance mechanosensitive channel**	**1.77**	**Cell wall/membrane/envelope biogenesis**	**55.79**	**11.12**	**20.17**	**20.24**	**3.22**	**5.09**
ERBA_00132	**MscS mechanosensitive ion channel**	**1.46**	**Cell wall/membrane/envelope biogenesis**	**99.66**	**24.29**	**37.39**	**28.19**	**10.58**	**10.58**
ERBA_02015	**ATP-grasp domain protein**	**1.37**	**Cell wall/membrane/envelope biogenesis**	**501.21**	**131.74**	**214.38**	**164.69**	**19.64**	**29.11**
ERBA_01271	**Lipoprotein-releasing system transmembrane protein LolC**	**1.29**	**Cell wall/membrane/envelope biogenesis**	**120.66**	**33.6**	**56.64**	**32.01**	**3.91**	**4.57**
ERBA_01460	**hypothetical protein**	**1.1**	**Cell wall/membrane/envelope biogenesis**	**182.72**	**57.66**	**47.12**	**47.15**	**12.35**	**7.48**
ERBA_01454	Immunogenic protein MPT70 analog	0.94	Cell wall/membrane/envelope biogenesis	188.66	67.44	40.39	30.5	18.37	5.02
ERBA_01860	UDP-4-amino-4-deoxy-L-arabinose--oxoglutarate aminotransferase	0.77	Cell wall/membrane/envelope biogenesis	147.89	58.8	154.17	53.98	9.41	6.6
ERBA_00275	L-carnitine-gamma-butyrobetaine antiporter	0.6	Cell wall/membrane/envelope biogenesis	182.67	82.43	152.98	45.95	2.28	12.19
ERBA_01449	**glycine N-methyltransferase**	**1.63**	**Coenzyme transport and metabolism**	**250.38**	**56.52**	**121.05**	**81.74**	**5.73**	**8.65**
ERBA_00049	**Geranylgeranyl diphosphate synthase**	**1.43**	**Coenzyme transport and metabolism**	**2530.78**	**636.9**	**1087.93**	**929.68**	**46.88**	**155.05**
ERBA_00074	**Biotin synthase**	**1.4**	**Coenzyme transport and metabolism**	**285.42**	**73.75**	**170.89**	**66.84**	**10.21**	**20.63**
ERBA_01620	**Digeranylgeranylglyceryl phosphate synthase**	**1.39**	**Coenzyme transport and metabolism**	**478.52**	**122.16**	**243.85**	**269.07**	**6.43**	**34.19**
ERBA_01927	**Phosphomethylpyrimidine synthase**	**1.38**	**Coenzyme transport and metabolism**	**237.87**	**61.84**	**120.2**	**103.18**	**11.12**	**20.93**
ERBA_00293	**MoeA region protein**	**1.34**	**Coenzyme transport and metabolism**	**195.2**	**51.74**	**131.93**	**97.81**	**9.25**	**15.11**
ERBA_00177	**Ubiquinone-menaquinone biosynthesis methyltransferase ubiE**	**1.27**	**Coenzyme transport and metabolism**	**191.18**	**53.3**	**128.28**	**72.49**	**14.58**	**20.85**
ERBA_01332	**Methylated-thiol--coenzyme M methyltransferase**	**1.22**	**Coenzyme transport and metabolism**	**74.14**	**21.44**	**32.66**	**38.85**	**5.07**	**6.19**
ERBA_02052	**prenyltransferase**	**1.17**	**Coenzyme transport and metabolism**	**119.4**	**36.29**	**94.69**	**23.36**	**1.98**	**17.54**
ERBA_00043	**Glutamyl-tRNA reductase**	**1.13**	**Coenzyme transport and metabolism**	**1536.67**	**476.74**	**805.15**	**546.14**	**22**	**145.09**
ERBA_01335	**phenylacetate--CoA ligase**	**1.1**	**Coenzyme transport and metabolism**	**60.98**	**19.64**	**49.13**	**9.25**	**2.68**	**6.21**
ERBA_00587	**putative molybdenum cofactor guanylyltransferase**	**1.1**	**Coenzyme transport and metabolism**	**117.84**	**37.5**	**77.99**	**46.05**	**8.02**	**25.12**
ERBA_00880	**Uroporphyrinogen-III decarboxylase Coenzyme metabolism**	**1.08**	**Coenzyme transport and metabolism**	**459.03**	**148.01**	**193.41**	**161.85**	**22.82**	**22.45**
ERBA_00042	**Delta-aminolevulinic acid dehydratase**	**1.05**	**Coenzyme transport and metabolism**	**1567.78**	**515.58**	**836.08**	**555.79**	**25.79**	**206.79**
ERBA_00044	Siroheme synthase	0.97	Coenzyme transport and metabolism	1045.91	361.97	443.55	384.3	26.23	74.68
ERBA_02111	Uroporphyrinogen-III synthase	0.94	Coenzyme transport and metabolism	474.44	168.76	282.33	140.67	11.74	40.04
ERBA_01491	methyltransferase	0.9	Coenzyme transport and metabolism	220.95	80.92	154.48	23.52	20.93	28.31
ERBA_00040	putative porphobilinogen deaminase	0.76	Coenzyme transport and metabolism	1976.49	794.81	1282.33	443.22	38.29	196.65
ERBA_01403	2--3'-dephosphocoenzyme-A synthase	0.75	Coenzyme transport and metabolism	190.04	77.75	232.93	6.56	3.82	21.71
ERBA_01930	Dihydropteroate synthase	0.71	Coenzyme transport and metabolism	872.65	362.92	409.61	269.33	27.86	23.07
ERBA_00041	Glutamate-1-semialdehyde 21-aminomutase	0.62	Coenzyme transport and metabolism	1853.98	825.4	1265.55	420.64	40.56	163.53
ERBA_00136	Siroheme synthase	0.58	Coenzyme transport and metabolism	333.06	153.86	275.09	20.4	10.56	22.84
ERBA_01201	Cobyrinic acid AC-diamide synthase	0.57	Coenzyme transport and metabolism	445.47	207.56	503.12	37.05	13.1	51.69
ERBA_00963	**Methylated-thiol--coenzyme M methyltransferase**	**1.95**	**Coenzyme transport and metabolism /General function prediction only**	**107.28**	**18.57**	**51.52**	**57.14**	**3.4**	**4.08**
ERBA_02171	**Nod factor export ATP-binding protein I**	**2.12**	**Defense mechanisms**	**302.77**	**47.38**	**104.3**	**87.77**	**3.06**	**27.19**
ERBA_01526	**Bacitracin transport ATP-binding protein BcrA**	**1.78**	**Defense mechanisms**	**145.95**	**28.31**	**80.33**	**80.54**	**12.21**	**4.71**
ERBA_01920	**Multidrug export protein MepA**	**1.39**	**Defense mechanisms**	**160.48**	**41.65**	**123.23**	**32.14**	**12.47**	**22.96**
ERBA_00654	Nod factor export ATP-binding protein I	0.86	Defense mechanisms	156.19	59.06	137.44	21.52	3.58	21.93
ERBA_00432	**Hydroxylamine reductase**	**2.4**	**Energy production and conversion**	**1328.38**	**176.1**	**1862.17**	**266.65**	**18.93**	**268.65**
ERBA_01549	**Nitrogen fixation protein NifU**	**1.71**	**Energy production and conversion**	**214.44**	**44.06**	**86.23**	**120.3**	**7.66**	**17.94**
ERBA_01103	**CoB--CoM heterodisulfide reductase 1 subunit B**	**1.63**	**Energy production and conversion**	**576.5**	**126.38**	**294.25**	**205.02**	**9.47**	**27.61**

ERBA_00577	**Putative fumarate hydratase subunit beta**	**1.6**	**Energy production and conversion**	**212.48**	**47.65**	**280.68**	**65.04**	**12.96**	**67.58**
ERBA_00903	**Tungsten-containing aldehyde ferredoxin oxidoreductase**	**1.58**	**Energy production and conversion**	**113.18**	**25.72**	**90.2**	**39.94**	**5.93**	**23.99**
ERBA_00976	**adenosine tRNA methylthiotransferase MiaB**	**1.57**	**Energy production and conversion**	**208.79**	**47.33**	**79.85**	**92.06**	**5.85**	**3.01**
ERBA_00576	**Putative fumarate hydratase subunit alpha**	**1.55**	**Energy production and conversion**	**343.14**	**79.47**	**320.21**	**78.79**	**8.87**	**43.32**
ERBA_00987	**geranylgeranyl reductase**	**1.52**	**Energy production and conversion**	**108.28**	**25.38**	**30.94**	**48.04**	**6.4**	**2.28**
ERBA_02172	**thioredoxin-related protein**	**1.47**	**Energy production and conversion**	**251.27**	**61.04**	**146.2**	**106.07**	**8.53**	**23.43**
ERBA_01878	**formylmethanofuran dehydrogenase subunit E**	**1.44**	**Energy production and conversion**	**373.17**	**93.4**	**87.9**	**163.33**	**13.92**	**6.38**
ERBA_01014	**adenosine tRNA methylthiotransferase MiaB**	**1.21**	**Energy production and conversion**	**366.31**	**107.29**	**276.13**	**120.4**	**10.52**	**29.16**
ERBA_02226	**Photosystem I iron-sulfur center**	**1.19**	**Energy production and conversion**	**557.92**	**165.99**	**324.53**	**134.96**	**34.89**	**49.53**
ERBA_00058	**V-type ATP synthase beta chain**	**1.16**	**Energy production and conversion**	**9118.51**	**2772.37**	**5073.48**	**2843.56**	**272.65**	**438.99**
ERBA_00260	**Acetyl-CoA decarbonylase-synthase complex subunit alpha**	**1.13**	**Energy production and conversion**	**854.04**	**263.74**	**793.28**	**468.44**	**26.04**	**264.92**
ERBA_02128	**Anaerobic nitric oxide reductase flavorubredoxin**	**1.12**	**Energy production and conversion**	**212.31**	**67.03**	**113.28**	**46.16**	**6.1**	**23.47**
ERBA_01164	**Cytochrome C oxidasecbb3-type protein**	**1.11**	**Energy production and conversion**	**91.48**	**28.6**	**49.06**	**29.59**	**8.83**	**12.28**
ERBA_00161	**Isopentenyl-diphosphate delta-isomerase**	**1.03**	**Energy production and conversion**	**450.95**	**150.64**	**290.28**	**159.57**	**19.29**	**64.83**
ERBA_01585	Phycocyanobilin lyase subunit alpha	0.97	Energy production and conversion	364.51	126.32	78.68	117.22	15.01	8.95
ERBA_01172	Plastocyanin	0.97	Energy production and conversion	219.99	77.07	119.58	18.39	7.63	18.88
ERBA_00259	Acetyl-CoA decarbonylase-synthase complex subunit epsilon	0.96	Energy production and conversion	654.94	227.13	581.78	357.2	38.1	179.98
ERBA_01132	CoB--CoM heterodisulfide reductase 1 iron-sulfursubunit A	0.96	Energy production and conversion	236.27	82.39	190.55	83.73	3.78	28.12
ERBA_00979	Pyridine nucleotide-disulfide oxidoreductase protein	0.96	Energy production and conversion	151.13	53.14	79.59	30.52	5.58	6.96
ERBA_01173	Plastocyanin	0.92	Energy production and conversion	495.43	177.81	272.09	98.88	41.78	29.63
ERBA_00790	Indolepyruvate oxidoreductase subunit IorA	0.91	Energy production and conversion	169.21	61.57	109.3	9.46	11.97	2.23
ERBA_02123	Flavohemoprotein	0.89	Energy production and conversion	109.33	40.09	76.31	14.08	15.52	20.99
ERBA_00202	Tungsten-containing aldehyde ferredoxin oxidoreductase	0.87	Energy production and conversion	108.81	40.62	89.14	20.02	1.72	7.82
ERBA_00057	V-type ATP synthase alpha chain	0.85	Energy production and conversion	6069.02	2293.64	3045.36	1985.25	198.69	236.26
ERBA_02075	putative succinate-semialdehyde dehydrogenase NADP	0.8	Energy production and conversion	251.6	99.09	134.56	56.46	8.83	34.47
ERBA_00646	Glutathione reductase	0.74	Energy production and conversion	181.49	74.21	142.38	31.19	18.59	37.6
ERBA_01481	Tungsten-containing formylmethanofuran dehydrogenase 2 subunit B	0.72	Energy production and conversion	373.03	154.77	298.49	80.93	8	24.61
ERBA_01482	Acetyl-CoA decarbonylase-synthase complex subunit gamma	0.71	Energy production and conversion	333.07	140.66	299.5	63.21	18.36	28.21
ERBA_00637	Nitroreductase Energy production	0.7	Energy production and conversion	632.45	266.42	377.21	137.47	7.99	25.85
ERBA_01069	Coenzyme F420-reducing hydrogenase	0.69	Energy production and conversion	262.36	111.89	131.17	12.87	14.36	31.41
ERBA_00059	V-type ATP synthase subunit D	0.68	Energy production and conversion	6840.96	2907.58	4740.47	2142.14	233.05	296.91
ERBA_00226	Carbon monoxide dehydrogenase 2	0.68	Energy production and conversion	226.75	97.23	155.21	42.76	2.98	17.58
ERBA_00013	radical SAM protein	0.63	Energy production and conversion	230.1	101.66	130.7	72.43	12.88	8.56
ERBA_00010	Glycolate oxidase subunit GlcD	0.54	Energy production and conversion	1399.49	661.03	715.35	197.57	25.37	48.68
ERBA_00011	CoB--CoM heterodisulfide reductase 2 iron-sulfursubunit D	0.49	Energy production and conversion	501.13	244.43	282.14	69.71	13.88	33.49
ERBA_00357	Oxaloacetate decarboxylase alpha chain	0.95	Energy production and conversion /Lipid transport and metabolism	331.89	117.77	240.12	60.35	4.7	56.52
ERBA_01547	**NAD-dependent dehydrogenase**	**3.11**	**Function unknown**	**93.12**	**6.92**	**21.29**	**35.6**	**5.68**	**14.1**
ERBA_00237	**Protein of unknown function DUF131**	**2.44**	**Function unknown**	**138.13**	**16.91**	**62.73**	**72.53**	**3.27**	**6.8**
ERBA_01404	**hypothetical protein**	**1.78**	**Function unknown**	**226.15**	**44.71**	**77.83**	**82.76**	**3.14**	**15.37**
ERBA_00674	**transporter**	**1.68**	**Function unknown**	**141.42**	**29.69**	**97.84**	**45.5**	**8.12**	**16.47**
ERBA_01058	**Thymidylate synthase protein**	**1.68**	**Function unknown**	**380.43**	**81.14**	**121.43**	**80.62**	**8.97**	**11.26**
ERBA_00128	**hypothetical protein**	**1.58**	**Function unknown**	**48.51**	**10.83**	**24.85**	**29.9**	**2.11**	**4.67**
ERBA_01527	**hypothetical protein**	**1.39**	**Function unknown**	**48.95**	**12.69**	**32.18**	**22.63**	**4.41**	**8.65**
ERBA_02008	**hypothetical protein**	**1.34**	**Function unknown**	**142.77**	**38.2**	**70.29**	**45.44**	**9.61**	**18.35**
ERBA_00881	**Inner membrane protein YaaH**	**1.33**	**Function unknown**	**712.59**	**193.95**	**172.84**	**208.88**	**6.28**	**26.78**

ERBA_00381	Phosphoenolpyruvate carboxylase	1.26	Function unknown	122.4	35	91.69	26.99	3.02	15.73
ERBA_01869	Protein of unknown function DUF99	1.25	Function unknown	449.06	128.47	179.8	96.03	11.82	7.35
ERBA_01755	DisA checkpoint controller nucleotide-binding protein	1.2	Function unknown	669.4	196.81	181.85	263.87	10.58	24.33
ERBA_00578	hypothetical protein	1.2	Function unknown	191.38	56.95	178.01	54.01	3.75	16.25
ERBA_01866	thymidylate synthase	1.19	Function unknown	192.94	57.95	111.74	45.76	13.91	8.38
ERBA_01015	hypothetical protein	1.19	Function unknown	162.82	48.76	76.5	49.67	10.57	11.27
ERBA_02011	IMPACT family member YvyE	1.17	Function unknown	95.58	29.36	69.06	26.28	13.87	14.7
ERBA_01707	Ribosomal protein S6 modification protein	1.13	Function unknown	341.15	106.73	241.02	133.39	11.23	42.83
ERBA_01493	hypothetical protein	1.09	Function unknown	56.38	17.91	38.63	28.56	2.83	2.9
ERBA_01926	GTP	1.07	Function unknown	201.31	65.23	90.8	63.54	14.18	7.56
ERBA_01187	hypothetical protein	1.06	Function unknown	150.72	50.25	79.12	21.48	11.22	25.33
ERBA_01009	Protease PrsW	1.05	Function unknown	71.14	23.31	42.69	18.23	3.68	3.95
ERBA_00029	hypothetical protein	1.03	Function unknown	726	243.81	468.15	83.98	13.61	58.4
ERBA_01470	hypothetical protein	1.01	Function unknown	82.72	27.64	83.1	28.41	9.61	1.96
ERBA_02127	rubrerythrin	0.99	Function unknown	288.05	99.77	262.16	12.72	24.59	16.04
ERBA_00133	hypothetical protein	0.93	Function unknown	113.41	41.3	41.65	14.46	14.37	4.8
ERBA_00271	Protein of unknown function DUF89	0.91	Function unknown	251.06	92.15	219.19	21.64	13.28	27.77
ERBA_00588	putative containing protein	0.89	Function unknown	182.73	67.11	181.67	45.93	8.83	25.15
ERBA_01718	Secreted effector protein pipB2	0.88	Function unknown	104.56	38.81	52.73	9.11	2.75	13.95
ERBA_02004	ApbE family lipoprotein	0.85	Function unknown	240.69	90.37	147.11	83.97	16.7	25.76
ERBA_00821	Periplasmic Fe hydrogenase large subunit	0.85	Function unknown	174.69	66.35	109.42	43.4	14.15	5.07
ERBA_00407	methanogenesis marker protein 12	0.84	Function unknown	144.55	55.03	98.3	16.02	12.52	23.77
ERBA_00392	methyltransferase	0.81	Function unknown	464.57	181.9	172.08	55.26	14.85	30.02
ERBA_00838	HTH DNA-binding protein	0.81	Function unknown	822.06	318.98	295.99	278.16	21.65	53.64
ERBA_01846	tRNA-splicing ligase RtcB	0.61	Function unknown	355.21	158.47	198.6	102.65	4.29	25.99
ERBA_00361	phosphoserine phosphatase	0.58	Function unknown	240.28	110.26	197.69	41.05	15.71	14.94
ERBA_02140	Archaeal of protein	0.57	Function unknown	413.11	190.57	336.33	70.79	12.33	23.43
ERBA_00964	Dimethylamine corrinoid protein 2	2.39	General function prediction only	65.86	8.12	31.14	45.71	3.11	13.7
ERBA_00169	tRNA -dimethyltransferase	1.88	General function prediction only	245.12	45.05	62.71	130.19	13.47	15.92
ERBA_00580	DNA protection protein DPS	1.8	General function prediction only	311.57	60.69	92.62	153.96	1.98	16.44
ERBA_01405	Protein TrpH	1.75	General function prediction only	953.77	191.59	274.12	341.7	22.7	9.89
ERBA_00807	2-succinyl-6-hydroxy-2 4-cyclohexadiene-1-carboxylate synthase	1.74	General function prediction only	449.28	91.92	251.02	59.08	19.43	7.98
ERBA_00528	Met-10+ related protein	1.73	General function prediction only	145.57	29.87	110.51	35.18	2.74	13.54
ERBA_00076	transcriptional regulator protein	1.68	General function prediction only	726.65	155.71	286	39.34	21.32	10.36
ERBA_01479	oxidoreductase molybdopterin binding domain containing protein	1.61	General function prediction only	131.88	28.97	45.09	52.66	16.27	11.48
ERBA_00647	PBP family phospholipid-binding protein	1.6	General function prediction only	105.45	23.58	52.7	37.99	9.63	13.45
ERBA_02110	iron-sulfur flavoprotein	1.59	General function prediction only	282.35	63.34	135.99	103.86	7.72	15.47
ERBA_00648	Uracil phosphoribosyltransferase	1.59	General function prediction only	50.29	11.06	25.15	23	4.59	8.12
ERBA_02068	Putative mycofactocin radical SAM maturase MftC	1.59	General function prediction only	78.14	17.61	43.82	24.45	5.02	7.06
ERBA_02094	Coenzyme PQQ synthesis protein E	1.53	General function prediction only	387.44	90.86	188.16	171.13	12.95	19.7
ERBA_00050	Coenzyme PQQ synthesis protein E	1.5	General function prediction only	2911.84	695.67	1230.84	1108.24	74.29	128.52
ERBA_01051	Dimethylamine corrinoid protein 2	1.48	General function prediction only	237.97	57.98	113.97	50.92	10.37	12.81
ERBA_00939	Phosphoglycolate phosphatase	1.41	General function prediction only	209.92	53.39	57.35	70.91	4.62	27.36
ERBA_01380	Archaea-specific pyridoxal phosphate-dependent enzyme General function prediction only	1.36	General function prediction only	385.79	102.99	308.78	75.16	16.22	36.79
ERBA_00393	NADPH-dependent FMN reductase protein	1.27	General function prediction only	176.74	50.06	95.34	23.56	16.48	20.83
ERBA_00368	Protein CbbY	1.26	General function prediction only	161.62	46.09	143.42	35.98	7.85	32.94
ERBA_01922	ABC transporter protein	1.25	General function prediction only	399.03	114.58	296.28	49.73	24.65	36.57
ERBA_00004	Putative NADH oxidase	1.2	General function prediction only	84.87	25.33	37.66	25.91	7.08	3.55
ERBA_00302	Putative mycofactocin radical SAM maturase MftC	1.2	General function prediction only	518.1	153.79	257.59	146.88	14.02	17.35
ERBA_01711	putative biotin transporter BioY	1.15	General function prediction only	107.41	32.58	42.26	52.78	4.81	10.56
ERBA_00047	Putative mycofactocin radical SAM maturase MftC	1.12	General function prediction only	1959.2	610.23	832.38	714.65	21.83	86.95
ERBA_01050	domain containing protein	1.12	General function prediction only	290.53	90.71	124.42	95.2	19.64	14.45
ERBA_00485	phosphoribosyltransferase	1.12	General function prediction only	141.63	44.93	64.12	6.8	6.66	24.12
ERBA_00904	NADPH-dependent FMN reductase protein	1.1	General function prediction only	144.43	46.42	167.68	13.63	5.77	30.99
ERBA_01990	Monomethylamine	1.08	General function	139.97	45.22	288.89	42.81	9.66	46.25

	corrinoid protein 1		**prediction only**				
ERBA_00403	**GMP-IMP nucleotidase YrfG**	**1.07**	**General function prediction only**	**157.73**	**51.35**	**104.5**	**16.49**	**17.84**	**38.03**
ERBA_00198	**hypothetical protein**	**1.01**	**General function prediction only**	**213.84**	**72.43**	**223.3**	**52.32**	**12.7**	**59.76**
ERBA_00062	Protein CbbQ	0.94	General function prediction only	1068.85	379.96	471.65	382.1	63.23	40.44
ERBA_00160	aspartate-glutamate-uridylate kinase	0.89	General function prediction only	315.92	116.25	203.31	77.01	5.87	28.43
ERBA_02122	NADH dehydrogenase	0.86	General function prediction only	109.99	41.1	72.94	19.94	11.05	12.92
ERBA_00158	Memo-like protein	0.81	General function prediction only	1552.73	608.82	1348.99	298.25	32.83	48.55
ERBA_01066	putative aspartoacylase	0.79	General function prediction only	102.92	40.75	54.22	34.43	7.31	10.61
ERBA_01716	permease	0.78	General function prediction only	173.13	68.9	67.96	45.95	5.21	18.36
ERBA_01758	Protein pelota	0.78	General function prediction only	121.02	48.55	98.63	26.3	12.46	25.03
ERBA_01152	Ribonuclease Z	0.75	General function prediction only	113.44	45.89	73.07	16.17	9.87	14.8
ERBA_00247	beta-lactamase-like protein	0.75	General function prediction only	779.4	316.07	879.89	186.8	3.53	32.86
ERBA_00411	putative manganese-dependent inorganic pyrophosphatase	0.74	General function prediction only	346.23	141.34	279.67	66.07	9.57	22.91
ERBA_00895	small GTP-binding protein	0.71	General function prediction only	167.89	70.88	186.14	10.95	5.93	25.47
ERBA_01757	NADPH-dependent FMN reductase protein	0.69	General function prediction only	201.33	85.37	110.58	19.15	15.82	7.91
ERBA_02148	Divalent metal cation transporter MntH	0.64	General function prediction only	641.24	282.32	340.76	130.29	31.02	19.77
ERBA_00538	metallophosphoesterase	0.61	General function prediction only	254.71	114.85	252.63	28.84	7.3	48.29
ERBA_01287	Succinoglycan biosynthesis transport protein ExoT	0.59	General function prediction only	121.87	55.41	98.59	23.33	8.2	7.48
ERBA_01592	**Dimethylamine corrinoid protein 1**	**1.67**	**General function prediction only / Coenzyme transport and metabolism**	**68.16**	**14.56**	**51.83**	**24.83**	**3.62**	**14.46**
ERBA_00850	Hydroxyacylglutathione hydrolase	0.76	General function prediction only /Inorganic ion transport and metabolism	121.25	48.84	92.87	24.01	4.74	22.53
ERBA_00822	**Sulfopyruvate decarboxylase subunit beta**	**1.07**	**General function prediction only /Multiple classes**	**204.51**	**65.88**	**139.4**	**89.35**	**8.84**	**7.25**
ERBA_02147	**Ferrous iron transport protein B**	**4.16**	**Inorganic ion transport and metabolism**	**4238.96**	**163.13**	**3907.05**	**252.13**	**9.95**	**279.99**
ERBA_00063	**Protein NorD**	**2.46**	**Inorganic ion transport and metabolism**	**379.35**	**46.23**	**101.27**	**202.22**	**2.96**	**16**
ERBA_01562	**putative calcium-transporting ATPase**	**2.44**	**Inorganic ion transport and metabolism**	**317.24**	**39.04**	**93.22**	**176.25**	**8.18**	**7.36**
ERBA_00064	**Arsenite resistance protein ArsB**	**2.24**	**Inorganic ion transport and metabolism**	**162.77**	**23.15**	**33.85**	**86.44**	**2.15**	**8.89**
ERBA_00581	**Zinc transporter ZupT**	**2.22**	**Inorganic ion transport and metabolism**	**123.59**	**17.89**	**56.36**	**29.32**	**7.29**	**17.68**
ERBA_02162	**cobalt ABC transporter ATP-binding protein**	**1.77**	**Inorganic ion transport and metabolism**	**521.84**	**103.38**	**185.23**	**218.82**	**13.09**	**11.42**
ERBA_01713	**membrane associated ATPase**	**1.51**	**Inorganic ion transport and metabolism**	**94.2**	**22.09**	**54.38**	**54.15**	**1.59**	**23.58**
ERBA_01107	**Fe3+-hydroxamate binding protein**	**1.49**	**Inorganic ion transport and metabolism**	**178.16**	**43.8**	**34.17**	**25.03**	**4.72**	**10.33**
ERBA_01147	**sodium-calcium exchanger protein**	**1.36**	**Inorganic ion transport and metabolism**	**58.95**	**15.54**	**22.31**	**15.99**	**2.55**	**1.13**
ERBA_01511	**Cobalamin import system permease protein BtuC**	**1.22**	**Inorganic ion transport and metabolism**	**87.36**	**25.08**	**78.5**	**46.5**	**2.37**	**15.98**
ERBA_01712	**Energy-coupling factor transporter ATP-binding protein EcfA**	**1.22**	**Inorganic ion transport and metabolism**	**134.62**	**38.58**	**47.27**	**78.13**	**3.25**	**6.56**
ERBA_01106	**FecCD transport protein**	**1.18**	**Inorganic ion transport and metabolism**	**91.75**	**27.75**	**22.76**	**32.2**	**5.4**	**8.21**
ERBA_02132	**Na antiporter**	**1.12**	**Inorganic ion transport and metabolism**	**48.85**	**15.33**	**24.88**	**17.66**	**4.07**	**11.68**
ERBA_01512	**ligand binding protein**	**1.07**	**Inorganic ion transport and metabolism**	**180.44**	**57.96**	**124.66**	**73.95**	**19.03**	**8.52**
ERBA_00093	**Trk system potassium uptake protein TrkA**	**1.03**	**Inorganic ion transport and metabolism**	**109.79**	**36.63**	**54.3**	**28.64**	**3.21**	**6.94**
ERBA_01297	Cadmium cobalt (cation efflux protein)	1	Inorganic ion transport and metabolism	92.1	31.18	59.1	36.03	3.58	9.03
ERBA_00094	Trk system potassium uptake protein TrkH	0.81	Inorganic ion transport and metabolism	78.19	30.36	42.27	23.63	1.09	6.82
ERBA_00095	Trk system potassium uptake protein TrkH	0.81	Inorganic ion transport and metabolism	92.56	35.76	53.43	25.03	10.61	8.99
ERBA_01064	Copper-exporting P-type ATPase A	0.79	Inorganic ion transport and metabolism	204.08	80.01	132.35	66.06	9.09	10.38
ERBA_01993	K+-dependent Na+-Ca+ exchanger related-protein	0.78	Inorganic ion transport and metabolism	91.66	36.34	64.16	21.31	8.06	7.92
ERBA_00140	Ferritin-1	0.65	Inorganic ion transport and metabolism	665.3	293.04	434.26	54.8	71.76	15.02
ERBA_00467	Signal recognition particle 54 kDa protein	0.79	Intracellular trafficking, secretion, and vesicular transport	570.72	225.97	637.11	133.24	25.67	57.93
ERBA_01822	Protein-export membrane protein SecF	0.61	Intracellular trafficking, secretion, and vesicular transport	483.9	217.09	473.48	83.69	9.24	30.78

ERBA_02260	Protein TolB	0.59	Intracellular trafficking, secretion, and vesicular transport	50.2	22.93	44.46	7.52	3.12	3.96
ERBA_00159	**Mevalonate kinase**	**1.12**	**Lipid transport and metabolism**	**833.12**	**261.94**	**656.37**	**195.89**	**21.35**	**67.87**
ERBA_01288	**Bifunctional protein GlmU**	**2.48**	**Multiple classes**	**131.92**	**15.61**	**26.32**	**84.96**	**4.65**	**9.07**
ERBA_02093	**HIT-like protein HinT**	**1.59**	**Multiple classes**	**399.61**	**88.91**	**196.22**	**199.1**	**22.87**	**59.22**
ERBA_00655	**Nodulation protein J**	**1.48**	**Multiple classes**	**255.94**	**61.42**	**128.43**	**144.46**	**9.36**	**17.33**
ERBA_01067	**Ribosomal protein S6 modification protein**	**1.39**	**Multiple classes**	**106.46**	**27.38**	**38.02**	**47.9**	**9.52**	**1.85**
ERBA_00649	**3-oxoacyl-acyl-carrier-protein reductase FabG**	**1.33**	**Multiple classes**	**57.61**	**15.72**	**36.17**	**13.33**	**2.96**	**10.18**
ERBA_00820	**Phosphoadenosine phosphosulfate reductase**	**1.27**	**Multiple classes**	**128.08**	**36.34**	**58.67**	**20.85**	**9.34**	**21.1**
ERBA_00038	putative dihydroorotate dehydrogenase B electrontransfer subunit	0.82	Multiple classes	807.93	314.58	565.29	119.82	52.51	98.8
ERBA_00414	putative brix domain-containing ribosomal biogenesis protein	0.81	Multiple classes	346.49	134.07	271.46	121.81	12.9	26.23
ERBA_02076	putative acetolactate synthase large subunit	0.8	Multiple classes	177.25	69.38	135.27	32.25	5.21	21.26
ERBA_01016	Bifunctional protein GlmU	0.68	Multiple classes	333.73	143.01	275.54	31.74	12.64	50.14
ERBA_00061	DEAD-box ATP-dependent RNA helicase CshA	0.65	Multiple classes	1668.81	729.69	950.24	274.17	83.5	45.53
ERBA_02144	Phosphoribosylformylglycinamidine synthase 2	0.99	Nucleotide transport and metabolism	1112.75	380.5	714.1	373.4	13.92	45.77
ERBA_00468	GMP synthase glutamine-hydrolyzing subunit A	0.84	Nucleotide transport and metabolism	249.75	95.41	266.53	57.15	7.53	38.46
ERBA_00039	Dihydroorotate dehydrogenase B catalytic subunit	0.79	Nucleotide transport and metabolism	1278.61	504.65	852.53	294.25	39.4	151.49
ERBA_00707	Adenylosuccinate lyase	0.53	Nucleotide transport and metabolism	205.89	98.1	233.29	41.44	10.97	20.13
ERBA_00563	**putative tRNA threonylcarbamoyladenosine biosynthesis protein KAE1**	**2.08**	**Posttranslational modification, protein turnover, chaperones**	**161.06**	**26.2**	**181**	**36.72**	**4.18**	**9.29**
ERBA_01209	**peptidase U32**	**1.99**	**Posttranslational modification, protein turnover, chaperones**	**714**	**120.29**	**220.5**	**383.17**	**24.54**	**42.8**
ERBA_01018	**FKBP-type peptidyl-prolyl cis-trans isomerase 2 Posttranslational modification**	**1.96**	**Posttranslational modification, protein turnover, chaperones**	**423.73**	**73.32**	**583.02**	**226.13**	**13.56**	**49.71**
ERBA_01175	**hypothetical protein**	**1.83**	**Posttranslational modification, protein turnover, chaperones**	**224.49**	**43.07**	**116.3**	**56.51**	**4.1**	**11.39**
ERBA_02173	**Glutaredoxin-1**	**1.74**	**Posttranslational modification, protein turnover, chaperones**	**123.96**	**25.02**	**42.17**	**27.42**	**8.6**	**18.14**
ERBA_01818	**Small heat shock protein C2**	**1.73**	**Posttranslational modification, protein turnover, chaperones**	**445.68**	**92.23**	**158.33**	**99.34**	**18.03**	**39.94**
ERBA_02229	**small heat shock protein**	**1.54**	**Posttranslational modification, protein turnover, chaperones**	**242.73**	**56.39**	**102.89**	**20.19**	**27.39**	**3.35**
ERBA_01819	**Small heat shock protein C2**	**1.47**	**Posttranslational modification, protein turnover, chaperones**	**575.1**	**142**	**260.87**	**146.26**	**15.09**	**29.96**
ERBA_00404	**Chaperone protein ClpB 2**	**1.43**	**Posttranslational modification, protein turnover, chaperones**	**321.98**	**81.86**	**284.97**	**64.38**	**7.03**	**71.9**
ERBA_00197	**Protein-L-isoaspartate O-methyltransferase 1**	**1.43**	**Posttranslational modification, protein turnover, chaperones**	**205.38**	**51.76**	**175.92**	**65.04**	**4.79**	**28.6**
ERBA_02227	**small heat shock protein**	**1.39**	**Posttranslational modification, protein turnover, chaperones**	**323.1**	**83.79**	**163.86**	**70**	**9.72**	**22.66**
ERBA_00747	**ADP-ribosylation-Crystallin J1**	**1.39**	**Posttranslational modification, protein turnover, chaperones**	**449.2**	**116.22**	**318.84**	**140.85**	**22.84**	**33.33**
ERBA_00986	**ATP-dependent 26S proteasome regulatory subunit Posttranslational modification**	**1.35**	**Posttranslational modification, protein turnover, chaperones**	**371.49**	**99.27**	**113.02**	**94.3**	**5.69**	**2.74**
ERBA_00825	**Archaeal Lon protease**	**1.24**	**Posttranslational modification, protein turnover, chaperones**	**1495.28**	**429.35**	**996.41**	**656.03**	**16.59**	**126.99**
ERBA_01817	**Small heat shock protein C2**	**1.23**	**Posttranslational modification, protein turnover, chaperones**	**281.78**	**81.66**	**142.65**	**62.68**	**25.65**	**37.68**
ERBA_02230	**small heat shock protein**	**1.19**	**Posttranslational modification, protein turnover, chaperones**	**247.08**	**73.76**	**96.59**	**80.42**	**30.85**	**15.9**
ERBA_00733	**peroxiredoxin 1**	**1.18**	**Posttranslational modification, protein turnover, chaperones**	**284.97**	**86.07**	**199.15**	**34.43**	**5.95**	**16.29**
ERBA_01908	**Intracellular alkaline protease**	**1.03**	**Posttranslational modification, protein turnover, chaperones**	**129.73**	**43.19**	**42.82**	**43.45**	**6.85**	**3**
ERBA_00091	**Chaperone protein DnaK**	**1.03**	**Posttranslational modification, protein turnover, chaperones**	**1604.8**	**542.49**	**1796.12**	**133.81**	**13.82**	**238.82**
ERBA_02160	10 kDa chaperonin	0.95	Posttranslational modification, protein	354.77	125.55	149.09	64.26	30.43	63.86

ERBA_02161	60 kDa chaperonin	0.92	Posttranslational modification, protein turnover, chaperones	1960.53	714.54	675.83	134.03	41.63	70.67
ERBA_01694	Proteasome-activating nucleotidase	0.9	Posttranslational modification, protein turnover, chaperones	326.16	120.24	457.02	22.36	10.42	22.72
ERBA_01052	cell division control protein 48	0.88	Posttranslational modification, protein turnover, chaperones	202.1	75.46	126.77	31.41	15.66	6.46
ERBA_00837	Protease HtpX	0.85	Posttranslational modification, protein turnover, chaperones	2355.88	901.27	989.92	130.4	46.17	23.89
ERBA_00809	Cytochrome c-type biogenesis protein CcmF	0.83	Posttranslational modification, protein turnover, chaperones	104.95	40.69	152.44	16.19	6.68	29.22
ERBA_01211	Cytochrome c biogenesis protein Posttranslational modification	0.8	Posttranslational modification, protein turnover, chaperones	178.79	70.77	51.41	22.08	7.28	8.41
ERBA_00048	Cytochrome c-type biogenesis protein CcmE	0.73	Posttranslational modification, protein turnover, chaperones	1013.06	415.33	605.68	334.75	22.19	20.38
ERBA_00191	peptidase U32	0.64	Posttranslational modification, protein turnover, chaperones	159.34	69.83	86.11	36.86	6.81	7.99
ERBA_01943	AIR synthase related protein	0.63	Posttranslational modification, protein turnover, chaperones	402.6	177.55	363.25	99.43	17.13	29.81
ERBA_00092	Chaperone protein DnaJ	0.62	Posttranslational modification, protein turnover, chaperones	994.17	448.48	1141.35	78.9	11.42	57.32
ERBA_01320	Pyruvate formate-lyase-activating enzyme	0.55	Posttranslational modification, protein turnover, chaperones	160.88	75.72	120.01	9.14	6.46	6.18
ERBA_02017	**transposase**	**2.29**	**Replication, recombination and repair**	**26.83**	**3.57**	**19.99**	**17.88**	**1.33**	**9.48**
ERBA_02158	**putative endonuclease 4**	**1.94**	**Replication, recombination and repair**	**526.25**	**93**	**192.03**	**184.93**	**6.22**	**16.79**
ERBA_00165	**transposase**	**1.76**	**Replication, recombination and repair**	**35.93**	**7.09**	**11.83**	**14.25**	**1.12**	**5.58**
ERBA_02330	**transposase**	**1.65**	**Replication, recombination and repair**	**62.45**	**13.67**	**32.42**	**23.64**	**5.21**	**6.11**
ERBA_00971	**DNA helicase**	**1.55**	**Replication, recombination and repair**	**57.92**	**13.46**	**28.35**	**28.31**	**3.65**	**7**
ERBA_02286	**transposase IS4 family protein**	**1.41**	**Replication, recombination and repair**	**175.22**	**44.61**	**104.27**	**56.41**	**7.61**	**7.44**
ERBA_00072	**Single-stranded-DNA-specific exonuclease RecJ**	**1.26**	**Replication, recombination and repair**	**418.8**	**119.38**	**381.46**	**92.88**	**1.72**	**46.75**
ERBA_01298	**DNA gyrase subunit B**	**1.19**	**Replication, recombination and repair**	**511.73**	**153.32**	**378.44**	**111.33**	**9.33**	**67.62**
ERBA_01295	**Type 2 DNA topoisomerase 6 subunit B**	**1.06**	**Replication, recombination and repair**	**317.69**	**103.06**	**317.68**	**113.45**	**11.78**	**17.24**
ERBA_00710	**UvrABC system protein A**	**1.05**	**Replication, recombination and repair**	**185.39**	**61.08**	**149.24**	**53.76**	**5.01**	**16.78**
ERBA_01838	UvrABC system protein C	0.96	Replication, recombination and repair	67.27	23.41	52.77	29.38	3.83	9.43
ERBA_00517	Transposase	0.91	Replication, recombination and repair	95.82	34.88	110.53	19.49	7.9	19.91
ERBA_01642	Exodeoxyribonuclease 7 large subunit	0.8	Replication, recombination and repair	402.37	157.17	180.5	92.45	10.94	21.13
ERBA_01326	transposase	0.77	Replication, recombination and repair	389.43	156.71	191.26	56.05	59.1	24.1
ERBA_01299	DNA gyrase subunit A	0.63	Replication, recombination and repair	368.9	163.18	370.12	66.66	16.67	50.18
ERBA_01389	**nemo like kinase**	**1.22**	**Replication, recombination and repair / Multiple classes**	**47.5**	**13.98**	**29.26**	**18.19**	**4.08**	**2.35**
ERBA_01059	**nemo like kinase**	**1.54**	**Replication, recombination and repair /Multiple classes**	**312.36**	**72.7**	**136.54**	**130.08**	**19.5**	**9.47**
ERBA_00003	RNA 3'-terminal phosphate cyclase	0.97	RNA processing and modification	305.89	106.6	125.88	75.9	7.35	8.13
ERBA_02016	**Imidazolonepropionase**	**2.02**	**Secondary metabolites biosynthesis, transport and catabolism**	**735.89**	**123.48**	**340.53**	**264.22**	**21.27**	**36.69**
ERBA_01213	**Virulence sensor protein BvgS**	**2.19**	**Signal transduction mechanisms**	**323.22**	**47.75**	**71.14**	**142.39**	**6.19**	**16.92**
ERBA_00065	**Universal stress protein YxiE**	**1.69**	**Signal transduction mechanisms**	**163.18**	**34.78**	**51.98**	**18.15**	**5.13**	**6.4**
ERBA_01002	**Blue-light-activated histidine kinase**	**1.67**	**Signal transduction mechanisms**	**32.29**	**6.7**	**21.58**	**15.98**	**2.69**	**0.55**
ERBA_00984	**Universal stress protein**	**1.53**	**Signal transduction mechanisms**	**99.96**	**23.73**	**57.39**	**9.83**	**9.19**	**14.14**
ERBA_01169	**Aerobic respiration control sensor protein ArcB**	**1.18**	**Signal transduction mechanisms**	**79.32**	**23.92**	**101.13**	**15.27**	**4.04**	**31.23**
ERBA_01583	**Serine-protein kinase RsbW**	**1.13**	**Signal transduction mechanisms**	**78.12**	**24.75**	**49.32**	**6.26**	**7.41**	**5.98**
ERBA_00829	**Phyllosphere-induced regulator PhyR**	**1.08**	**Signal transduction mechanisms**	**219.97**	**71.04**	**105.76**	**13.77**	**9.87**	**22.95**
ERBA_01145	**KaiC**	**1.06**	**Signal transduction mechanisms**	**237.71**	**77.93**	**196.46**	**56.5**	**6.25**	**11.68**
ERBA_02131	**Inosine-5'-monophosphate dehydrogenase**	**1.02**	**Signal transduction mechanisms**	**117.13**	**39.23**	**42.48**	**24.64**	**9.91**	**14.06**
ERBA_01983	Blue-light-activated histidine kinase	0.95	Signal transduction mechanisms	86.44	30.39	51.04	29.87	1.6	3.2

ERBA_00590	RIO-type serine-threonine-protein kinase Rio2	0.92	Signal transduction mechanisms	193.68	69.49	103.28	55.57	7.41	7.04
ERBA_01386	C4-dicarboxylate transport sensor protein DctS	0.91	Signal transduction mechanisms	68.84	25.16	74.08	12.46	5.28	7.38
ERBA_00394	PAS-PAC sensor signal transduction histidine kinase	0.85	Signal transduction mechanisms	65.39	25.04	23.63	8.54	2.28	5.04
ERBA_01462	Virulence sensor protein BvgS	0.69	Signal transduction mechanisms	106.56	45.48	89.7	7.66	6.04	15.77
ERBA_01628	Blue-light-activated histidine kinase	0.69	Signal transduction mechanisms	118.3	49.99	121.06	45.71	3.29	10
ERBA_01176	Blue-light-activated histidine kinase	0.56	Signal transduction mechanisms	97.26	45.7	77.67	4.09	4.67	12.77
ERBA_01104	**Transcriptional regulator MntR**	**2.86**	**Transcription**	**2188.97**	**207.1**	**1072.47**	**421.03**	**13.39**	**237.19**
ERBA_01500	**transcriptional regulator**	**2.59**	**Transcription**	**318.58**	**36.41**	**21.02**	**95.46**	**19.73**	**17.6**
ERBA_00433	**HxlR family transcriptional regulator**	**2.18**	**Transcription**	**1106.65**	**168.72**	**577.78**	**246.52**	**45**	**81.29**
ERBA_01226	**Transcription initiation factor IIB**	**1.77**	**Transcription**	**907.35**	**181.52**	**262.03**	**287.3**	**17.8**	**42.97**
ERBA_01571	**Winged helix-turn-helix protein**	**1.3**	**Transcription**	**1559.67**	**437.57**	**532.32**	**61.86**	**57.73**	**120.41**
ERBA_00046	**Protein NirD**	**1.13**	**Transcription**	**1902.87**	**590.4**	**878.61**	**722.83**	**53.21**	**155.69**
ERBA_00688	**DNA-directed RNA polymerase subunit B"**	**1.11**	**Transcription**	**1270.77**	**399.88**	**841.95**	**424.71**	**33.96**	**53.21**
ERBA_00045	**Protein NirD**	**1.07**	**Transcription**	**1432.27**	**463.01**	**622.17**	**471.79**	**33.11**	**56.71**
ERBA_00167	Transcription initiation factor IIB	0.93	Transcription	1855.99	663.8	646.53	420.75	68.75	27.7
ERBA_02007	transcriptional regulator protein	0.81	Transcription	1128.86	447.93	334.38	165.66	52.12	96.17
ERBA_01654	transcriptional regulator TrmB	0.74	Transcription	149.97	61.54	76.19	11.65	17.1	9.46
ERBA_02145	**Iron-dependent repressor IdeR**	**3.39**	**Transcription /Inorganic ion transport and metabolism**	**2612.08**	**171.31**	**2594**	**285.18**	**12.44**	**123.28**
ERBA_01402	**putative tRNA-dihydrouridine synthase**	**1.2**	**Transcription /Translation, ribosomal structure and biogenesis**	**447.99**	**134.49**	**634.46**	**1.99**	**7.07**	**36.62**
ERBA_01621	**tRNA-guanine transglycosylase**	**2.24**	**Translation, ribosomal structure and biogenesis**	**1278.28**	**183.29**	**463.55**	**572.7**	**5.26**	**46.31**
ERBA_01770	**Ribosomal RNA small subunit methyltransferase F**	**2.17**	**Translation, ribosomal structure and biogenesis**	**49.85**	**7.33**	**37.64**	**31.35**	**0.9**	**2.61**
ERBA_01070	**Aspartyl-glutamyl-tRNA amidotransferase subunit C**	**1.2**	**Translation, ribosomal structure and biogenesis**	**195.51**	**57.59**	**121.52**	**40.67**	**16.92**	**26.73**
ERBA_02335	**queuine tRNA-ribosyltransferase containing PUA domain protein**	**1.17**	**Translation, ribosomal structure and biogenesis**	**150.1**	**45.51**	**98.68**	**43.84**	**7.88**	**14.81**
ERBA_01071	**Glutamyl-tRNA amidotransferase subunit A**	**1**	**Translation, ribosomal structure and biogenesis**	**257.68**	**88.41**	**154.41**	**20.86**	**5.04**	**10.06**
ERBA_00300	50S ribosomal protein L10e	0.98	Translation, ribosomal structure and biogenesis	1769.49	623.09	1468.59	194.49	42.52	95.96
ERBA_00030	Glutamate--tRNA ligase	0.93	Translation, ribosomal structure and biogenesis	929.52	334.54	576.25	160.94	19.59	68.54
ERBA_00788	50S ribosomal protein L3	0.89	Translation, ribosomal structure and biogenesis	1711.36	635.07	639.46	297.32	8.09	69.53
ERBA_01600	Glycine--tRNA ligase	0.84	Translation, ribosomal structure and biogenesis	178.78	68.59	237	26	2.34	31.53
ERBA_00303	Tryptophan--tRNA ligase	0.79	Translation, ribosomal structure and biogenesis	327.96	129.65	243.68	88.45	5.89	40.71
ERBA_00787	50S ribosomal protein L4	0.75	Translation, ribosomal structure and biogenesis	1880.17	764.5	743.77	425.57	21.95	152.96
ERBA_02070	Ribose 15-bisphosphate isomerase	0.74	Translation, ribosomal structure and biogenesis	196.12	80.39	131.53	48.98	9.32	20.96
ERBA_01606	haloacid dehalogenase	0.71	Translation, ribosomal structure and biogenesis	300.2	125.37	190.2	57.15	13.33	5.79
ERBA_00786	50S ribosomal protein L23P	0.67	Translation, ribosomal structure and biogenesis	1224.67	529.64	471.53	295.48	80.19	50.27
ERBA_00187	Lysine--tRNA ligase 1	0.65	Translation, ribosomal structure and biogenesis	252.87	111.28	295.63	23.25	8.81	20.44
ERBA_00545	hypothetical protein	0.63	Translation, ribosomal structure and biogenesis	835.27	372.6	871	45.97	42.79	62.19
ERBA_00157	30S ribosomal protein S2	0.58	Translation, ribosomal structure and biogenesis	1537.2	706.06	1815	190.21	94.71	86.28
ERBA_01924	mRNA 3-end processing factor	0.54	Translation, ribosomal structure and biogenesis	908	429.48	867.47	93.27	12.79	45.96
ERBA_01072	Aspartyl-glutamyl-tRNA amidotransferase subunit B	0.53	Translation, ribosomal structure and biogenesis	252.13	120.23	224.69	27.85	9.86	5.72
ERBA_00926	Proline--tRNA ligase	0.51	Translation, ribosomal structure and biogenesis	474.26	228.61	370.6	27.57	26.39	18.51
ERBA_01421	Translation initiation factor 2 subunit gamma	0.51	Translation, ribosomal structure and biogenesis	551.43	266.51	801.53	53.82	11.55	47.18
ERBA_00141	Valine--tRNA ligase	0.51	Translation, ribosomal structure and biogenesis	493.79	238.95	497.31	30.11	16.79	31.97
ERBA_00142	Pyrrolysine--tRNA ligase	0.5	Translation, ribosomal structure and biogenesis	412.04	198.99	405.88	80.83	15.31	22.49
ERBA_00732	**Ribosomal RNA large subunit methyltransferase E**	**1.16**	**Translation, ribosomal structure and biogenesis /General function prediction only**	**947.46**	**290.07**	**842.08**	**224.85**	**4.22**	**72.5**
ERBA_00301	Translation initiation factor 2 subunit beta	0.88	Translation, ribosomal structure and biogenesis /General function prediction only	1419.72	535.72	1038.98	72.22	41.54	28.44

ERBA_02249	hypothetical protein	6.28	184.06	1.16	17.72	145.02	2.02	18.55
ERBA_02250	hypothetical protein	4.86	110.9	1.91	8.22	86.52	3.31	9.74
ERBA_02247	hypothetical protein	4.45	120.28	2.77	0	97.44	4.8	0
ERBA_00176	hypothetical protein	4.39	169.82	4.56	44.27	101.61	7.89	12.2
ERBA_02248	hypothetical protein	4.31	217.01	7.12	26.41	144.24	5.45	3.76
ERBA_01708	hypothetical protein	3.98	58.12	1.87	17.05	29.64	3.24	14.44
ERBA_02246	hypothetical protein	3.86	462.71	21.1	62.17	319.78	1.97	11.58
ERBA_02146	hypothetical protein	3.81	1084.46	51.84	1083.06	310.13	12.63	31.47
ERBA_02251	hypothetical protein	3.7	105.87	5.16	12.54	84.61	2.63	12.73
ERBA_02252	hypothetical protein	3.61	82.93	4.38	20.53	56.98	2.41	9.33
ERBA_01497	chloride channel core	3.57	85.6	4.65	10.21	37.2	8.05	9.32
ERBA_02092	transposase OrfB family protein	3.55	298.76	16.49	124.6	166.23	3.74	35.87
ERBA_02241	hypothetical protein	3.45	572.62	33.87	98.1	206.44	27.08	15.84
ERBA_01495	putative ATP-dependent Na+ efflux pump	3.31	89.79	5.36	2.65	4.39	9.28	4.58
ERBA_00401	hypothetical protein	3.01	54.03	3.42	3.45	30.12	5.92	5.98
ERBA_00078	transposase IS4 family protein	2.83	172.9	16.13	56.66	53.54	8.07	21.82
ERBA_02254	hypothetical protein	2.66	27.36	2.67	21.65	17.32	2.33	10.04
ERBA_01494	hypothetical protein	2.37	148.59	19.32	28.19	42.77	0.75	14.57
ERBA_02274	putative CopG-Arc-MetJ family transcriptional regulator	2.33	165.06	21.68	45.87	134.59	2.92	10.73
ERBA_02309	hypothetical protein	2.27	447.32	62.04	254.12	275.91	14.11	41.12
ERBA_01501	hypothetical protein	2.25	392.13	57.21	37.46	109.77	13.21	8.4
ERBA_01496	hypothetical protein	2.21	141.47	20.46	18.83	28.36	6.89	8.74
ERBA_02272	hypothetical protein	2.16	175.86	25.88	65.28	128.95	7.71	44.44
ERBA_00077	transposase IS4	2.13	344.47	53.44	91.31	73.14	8.68	16.23
ERBA_00905	hypothetical protein	2.04	210.14	35.42	78.94	90.84	5.77	18.66
ERBA_02253	Carboxypeptidase regulatory-like domain protein	2.03	47.71	7.94	23.49	28.45	5.48	7.29
ERBA_01294	Plasmid pRiA4b ORF-3-like protein	2.01	228.94	38.29	48.71	85.42	14	11.26
ERBA_01499	hypothetical protein	1.93	159.29	28.33	21.09	43.91	4.61	8.2
ERBA_01570	Tricorn protease-interacting factor F3	1.92	504.4	90.33	72	223.43	15.18	11.62
ERBA_01342	hypothetical protein	1.9	34.25	6.19	17.84	17.96	2.89	1.34
ERBA_00579	Protoglobin domain containing protein	1.88	116.61	21.3	67.05	32.22	8.38	28.89
ERBA_00941	transmembrane protein MttP	1.88	205.42	37.53	140.29	107.65	6.43	24.64
ERBA_01206	hypothetical protein	1.86	50.15	9.11	24.97	21.46	3.84	8.42
ERBA_01923	hypothetical protein	1.84	563.61	106.45	545.8	146.47	15.91	28.16
ERBA_00402	trimethylamine permease	1.83	146.27	27.96	96.01	19.53	7.21	15.12
ERBA_00928	hypothetical protein	1.77	237.85	46.61	89.26	127.15	4.65	42.43
ERBA_01625	monomethylamine permease	1.76	96.03	18.99	101.33	31.86	5.11	47.42
ERBA_01584	hypothetical protein	1.74	98.58	19.78	106.72	33.49	8.14	33.77
ERBA_02257	hypothetical protein	1.71	53.86	11.28	12.14	12.73	7.86	6.31
ERBA_00560	Serine-aspartate repeat-containing protein F	1.71	172.68	35.64	59.22	71.1	12.66	20.67
ERBA_02159	Winged helix-turn-helix DNA-binding protein	1.69	810.91	170.98	124.68	284.7	8.56	17.53
ERBA_02308	hypothetical protein	1.68	562.79	119.1	375.19	146.02	16.85	78.43
ERBA_01828	hypothetical protein	1.67	77.47	16.4	41.26	33.28	11.76	4
ERBA_00367	S-layer-related protein	1.66	468.88	100.73	142.63	153.49	16.95	29.85
ERBA_01670	hypothetical protein	1.66	373.25	81.25	33.53	23.6	21.08	34.04
ERBA_02001	hypothetical protein	1.65	51.44	10.92	23.55	24.13	6.91	7.18
ERBA_02170	Nodulation protein J	1.65	161.52	34.98	77.6	56.5	2.34	4.31
ERBA_01490	Citrate-proton symporter	1.64	86.23	18.28	36.68	44.19	13.42	7.18
ERBA_00012	hypothetical protein	1.64	114.54	25	44.56	32.29	3.05	5.28
ERBA_02051	hypothetical protein	1.62	938.46	209.29	1088.35	136.87	31.44	163.66
ERBA_01017	nucleic acid binding OB-fold tRNA-helicase-type	1.55	105.82	24.27	205.98	52.91	8.99	51.18
ERBA_02121	Cytochrome	1.54	270.4	63.47	110.37	61.99	6.73	4.81
ERBA_02313	glycosyltransferase	1.44	593.35	149.16	174.86	179.54	4.57	18.98
ERBA_01626	hypothetical protein	1.36	169.05	44.87	136.22	24.89	2.77	28.28
ERBA_01290	AbrB family transcriptional regulator	1.35	118.04	31.22	61.46	46.78	14.27	12.11
ERBA_00036	hypothetical protein	1.35	195.9	52.06	82.55	80.3	5.63	40.46
ERBA_02228	small heat shock protein	1.34	173.95	46.42	73.95	36.79	10.71	27
ERBA_01079	hypothetical protein	1.32	207.43	57.29	151.5	53.19	10.39	18.84
ERBA_02258	integrase family protein	1.32	329.18	90.31	77.21	103.74	16.52	14.16
ERBA_02326	hypothetical protein	1.32	106.5	28.99	41.09	39.63	2.73	11.05
ERBA_01498	hypothetical protein	1.29	308.54	86.34	63.29	95.83	5.5	15.94
ERBA_02270	hypothetical protein	1.29	201.38	56.56	96.05	64.8	30.01	28.62
ERBA_02244	hypothetical protein	1.28	437.5	121.49	127.22	197.87	16.4	33.4
ERBA_01572	hypothetical protein	1.28	230.5	64.07	72.46	58.33	16.78	23.37
ERBA_01457	hypothetical protein	1.23	125.43	36.93	132.41	25.81	7.11	23.83
ERBA_01573	Transmembrane exosortase protein	1.2	280.3	83.4	129.66	77.83	9.86	10.2
ERBA_00938	hypothetical protein	1.2	320.54	95.24	79.53	40.61	16.92	32.88
ERBA_01244	hypothetical protein	1.19	156.46	46.72	113.47	35.99	2.55	15.71
ERBA_00446	hypothetical protein	1.19	107.14	32.09	45.18	18.84	3.66	11.59
ERBA_02108	nigerythrin	1.16	399.15	121.55	175.54	132.37	17.9	32
ERBA_01574	hypothetical protein	1.15	513.34	156.8	183.2	179.86	24.42	9.27
ERBA_00639	family 2 glycosyl transferase	1.13	92.63	28.86	45.13	24.68	2.75	8.67
ERBA_01579	hypothetical protein	1.12	687.45	211.3	187.87	429.77	47.5	14.13
ERBA_00586	Calcium-dependent protease	1.08	92.07	29.4	76.17	42.89	5.1	14.39

ERBA_01148	**TM helix protein**	**1.07**	**63.89**	**20.5**	**68.96**	**24.73**	**9.93**	**14.54**
ERBA_01171	**hypothetical protein**	**1.06**	**125.7**	**41.34**	**58.95**	**22.88**	**6.83**	**19.2**
ERBA_02245	**hypothetical protein**	**1.04**	**614.96**	**200.77**	**228.52**	**315.53**	**5.2**	**32.86**
ERBA_02342	methanol protein	0.96	320.35	112.83	204.23	102.01	6.49	15.1
ERBA_01214	hypothetical protein	0.94	181.66	64.19	120.78	35.76	20.84	35.76
ERBA_02343	Methylated-thiol--corrinoid protein MtsB	0.93	1276.53	461.31	285.74	149.17	22.45	51.84
ERBA_00836	hypothetical protein	0.92	355.03	129.31	98.43	41.68	8.84	27.06
ERBA_00028	hypothetical protein	0.89	482.43	181.1	378.08	31.66	55.21	106.19
ERBA_01929	zinc-finger protein	0.87	524.07	195.03	221.85	139.41	4.3	24.07
ERBA_01596	hypothetical protein	0.84	248.11	94.39	172.43	35.85	26.04	16.03
ERBA_01309	hypothetical protein	0.75	823.21	338.31	903.72	98.05	27.73	37.6
ERBA_02269	hypothetical protein	0.74	526.7	215.1	284.88	106.55	34.31	7.28
ERBA_00360	S-layer protein	0.69	215.32	92.79	144.19	8.52	20.25	8.63
ERBA_01639	hypothetical protein	0.67	594.96	255.81	287.49	153.71	22.32	53.78
ERBA_00882	methanol protein	0.6	1318.86	598.12	444.76	59.45	55.3	51.23
ERBA_00234	hypothetical protein	0.6	2861.91	1292.7	1042.97	633.27	52.95	81.35
ERBA_01559	hypothetical protein	0.58	418.06	193.08	313.22	33.07	16.04	40.2
ERBA_01392	CoB--CoM heterodisulfide reductase 2 subunit E	0.54	1443.19	681.15	1399.68	141.83	28.03	32.59

Table S7. 360 down-regulated genes at 5% salinity relative to 10% salinity. (Genes with log2 fold change less than -1 are in bold.)

Gene ID	COG classification	log2 fold change	INDIGO annotation	FPKM avg. 5% NaCl	FPKM avg. 10% NaCl	FPKM avg. 15% NaCl	std.dev for 5% NaCl	std.dev for 10% NaCl	std.dev for 15% NaCl
ERBA_01659	**Amino acid transport and metabolism**	**-2.42**	**Molybdate-tungstate import ATP-binding protein WtpC**	**216.78**	**794.15**	**661.44**	**49.77**	**60.09**	**52.09**
ERBA_00743	**Amino acid transport and metabolism**	**-1.43**	**Glycine betaine-binding periplasmic protein**	**1618.11**	**3077.17**	**3546.53**	**551.92**	**372.15**	**479.14**
ERBA_01790	**Amino acid transport and metabolism**	**-1.3**	**Anthranilate phosphoribosyltransferase**	**251.95**	**443.74**	**438.04**	**152.1**	**118.2**	**55.86**
ERBA_01789	**Amino acid transport and metabolism**	**-1.22**	**N-anthranilate isomerase 1**	**117.4**	**192.81**	**153.03**	**46.72**	**27.71**	**12.33**
ERBA_00906	**Amino acid transport and metabolism**	**-1.09**	**ABC glycine betaine-L-proline transporter substrate-binding subunit**	**1643.56**	**2477.78**	**880.51**	**747.91**	**68.49**	**79.99**
ERBA_00616	**Amino acid transport and metabolism**	**-1.08**	**Glutathione-binding protein GsiB**	**85.83**	**126.16**	**65.34**	**9.42**	**14.27**	**9.86**
ERBA_00742	**Amino acid transport and metabolism**	**-1.05**	**Glycine betaine-binding periplasmic protein**	**2694.88**	**3932.83**	**3185.16**	**911.01**	**132.17**	**152.93**
ERBA_01085	Amino acid transport and metabolism	-0.79	Glutamate synthase large subunit-like protein	250.74	300.68	190.98	30.16	13.61	7.14
ERBA_00453	Amino acid transport and metabolism	-0.7	4-hydroxy-tetrahydrodipicolinate reductase	337.56	380.61	541.61	32.09	38.61	80.52
ERBA_00016	Amino acid transport and metabolism	-0.55	putative 2-isopropylmalate synthase	905.22	921.16	921.84	98.7	25.85	80.7
ERBA_01814	Amino acid transport and metabolism	-0.52	Aspartokinase	178.82	178.38	339.99	24.63	9.12	11
ERBA_00272	**Carbohydrate transport and metabolism**	**-1.13**	**23-bisphosphoglycerate-independent phosphoglycerate mutase**	**85.46**	**129.54**	**89.45**	**10.92**	**7.52**	**15.52**
ERBA_00495	Carbohydrate transport and metabolism	-0.68	Ribose-5-phosphate isomerase A	127.95	141.99	250.21	18.65	19.27	8.1
ERBA_00268	Carbohydrate transport and metabolism	-0.66	Glyceraldehyde-3-phosphate dehydrogenase	396.16	433.5	630.26	45.33	14.13	29.63
ERBA_01250	Carbohydrate transport and metabolism	-0.61	glucoamylase	91.05	95.8	107.73	16.15	8.66	7.06
ERBA_01273	Carbohydrate transport and metabolism	-0.51	Enolase	274.52	270.56	366.46	31.29	17.39	30.56
ERBA_02218	**Cell cycle control, cell division, chromosome partitioning**	**-3.05**	**chromosome partitioning ATPase protein**	**211.12**	**1218.58**	**93.78**	**24.73**	**21.47**	**5.54**
ERBA_02104	Cell cycle control, cell division, chromosome partitioning	-0.99	Chromosome partition protein Smc	170.23	231.92	179.69	30.74	16.38	16.49
ERBA_01043	Cell cycle control, cell division, chromosome partitioning	-0.78	Cell division protein FtsZ 3	373.57	442.96	358.53	38.8	21.92	16.11
ERBA_00505	Cell cycle control, cell division, chromosome partitioning	-0.75	Septum site-determining protein MinD	384.69	451.01	447.48	74.63	22.41	74.88
ERBA_01566	Cell cycle control, cell division, chromosome partitioning	-0.72	Cell division protein FtsZ 2	1162.58	1343.08	1022.07	298.95	41.66	111.65
ERBA_00607	**Cell motility**	**-4.36**	**Flagellin B2**	**1400.6**	**20573.44**	**4895.24**	**1002.33**	**1149.95**	**568.33**
ERBA_00287	**Cell motility**	**-1.63**	**flagella protein**	**283.89**	**601.55**	**281.69**	**64.89**	**61.16**	**23.45**
ERBA_01239	**Cell motility**	**-1.18**	**flagellar protein F**	**209.77**	**331.6**	**135.26**	**40.9**	**33.82**	**8.09**
ERBA_01240	Cell motility	-0.98	flagella protein	349.05	480.35	217.46	69.34	17.17	29.81
ERBA_01238	Cell motility	-0.85	flagellar protein G	145.89	182.71	69.75	31.13	23.3	9.65
ERBA_01281	**Cell wall/membrane/envelope biogenesis**	**-1.47**	**Glycosyltransferase Cell envelope biogenesi**	**56.19**	**108.57**	**96.31**	**12.49**	**9.2**	**22.94**
ERBA_00596	**Cell wall/membrane/envelope biogenesis**	**-1.47**	**N-acetyltransferase GCN5**	**33.61**	**63.85**	**33.21**	**9.17**	**17**	**5.25**
ERBA_01280	Cell wall/membrane/envelope biogenesis	-0.94	Undecaprenyl-phosphate 4-deoxy-4-formamido-L-arabinose transferase	173.26	230.67	173.47	10.79	43.38	23.88
ERBA_01278	Cell wall/membrane/envelope biogenesis	-0.81	glycoprotein 3-alpha-L-fucosyltransferase	153.81	186.2	184.02	17.18	31.73	39.99
ERBA_01656	Cell wall/membrane/envelope biogenesis	-0.79	hypothetical protein	431.77	511.86	418.33	87.71	52.94	41.55
ERBA_01657	Cell wall/membrane/envelope biogenesis	-0.58	hypothetical protein	529.45	542.59	454.56	81.27	23.82	50.34
ERBA_01740	**Chromatin structure and dynamics**	**-1.54**	**Archaeal histone A**	**248.91**	**505.9**	**354.78**	**81.5**	**51.54**	**76.73**

ERBA_01734	**Coenzyme transport and metabolism**	**-2.47**	**Trimethylamine methyltransferase MttB**	**4636.38**	**18026.02**	**4963.65**	**1519.12**	**1922.67**	**537.4**
ERBA_00843	**Coenzyme transport and metabolism**	**-2.38**	**Methyl-coenzyme M reductase I operon protein C**	**1283.62**	**4648.87**	**1202.4**	**288.7**	**432.34**	**289.36**
ERBA_00842	**Coenzyme transport and metabolism**	**-2.19**	**Methyl-coenzyme M reductase subunit gamma**	**1624.57**	**5171.23**	**1452.08**	**395.84**	**457.92**	**270.18**
ERBA_00844	**Coenzyme transport and metabolism**	**-2.02**	**Methyl-coenzyme M reductase II operon protein D**	**1797.63**	**5083.93**	**1837.19**	**469.48**	**479.22**	**243.11**
ERBA_00845	**Coenzyme transport and metabolism**	**-1.85**	**Methyl-coenzyme M reductase II subunit beta**	**2256.57**	**5665.47**	**2374.78**	**392.18**	**430**	**336.08**
ERBA_00841	**Coenzyme transport and metabolism**	**-1.73**	**Methyl-coenzyme M reductase II subunit alpha**	**2927.67**	**6754.37**	**3183.01**	**580.41**	**635.4**	**304.46**
ERBA_00143	**Coenzyme transport and metabolism**	**-1.72**	**Uroporphyrinogen-III decarboxylase Coenzyme metabolism**	**3866.56**	**9048.37**	**4478.16**	**1906.06**	**745.08**	**877**
ERBA_01809	**Coenzyme transport and metabolism**	**-1.58**	**Tetrahydromethanopterin S-methyltransferase subunit G**	**405.05**	**853.88**	**841.9**	**171.71**	**200.89**	**26.11**
ERBA_01808	**Coenzyme transport and metabolism**	**-1.51**	**Tetrahydromethanopterin S-methyltransferase subunit F**	**480.74**	**960.21**	**915.88**	**132.32**	**99.86**	**84.87**
ERBA_01003	**Coenzyme transport and metabolism**	**-1.32**	**Uroporphyrinogen-III decarboxylase Coenzyme metabolism**	**60.51**	**103.89**	**124.73**	**21.5**	**8.43**	**23.2**
ERBA_00871	**Coenzyme transport and metabolism**	**-1.26**	**Adenosylhomocysteinase**	**4026.17**	**6718.35**	**3340.25**	**734.97**	**649.4**	**317.25**
ERBA_01807	**Coenzyme transport and metabolism**	**-1.18**	**Tetrahydromethanopterin S-methyltransferase subunit A**	**981.89**	**1557.34**	**1631.2**	**197.83**	**102.92**	**99.03**
ERBA_00872	**Coenzyme transport and metabolism**	**-1.17**	**S-adenosylmethionine synthase**	**3541.53**	**5601**	**3131.66**	**1100.58**	**580**	**362.01**
ERBA_00281	**Coenzyme transport and metabolism**	**-1.11**	**von Willebrand factor type A**	**60.69**	**89.58**	**47.78**	**15.4**	**4.28**	**2.13**
ERBA_00615	**Coenzyme transport and metabolism**	**-1.08**	**Ubiquinone-menaquinone biosynthesis methyltransferase ubiE**	**61.57**	**89.96**	**48.62**	**18.22**	**28.03**	**1.91**
ERBA_01810	**Coenzyme transport and metabolism**	**-1.06**	**Tetrahydromethanopterin S-methyltransferase subunit H**	**1193.73**	**1739.94**	**1887.71**	**291.74**	**169.84**	**58.16**
ERBA_01806	Coenzyme transport and metabolism	-0.99	Tetrahydromethanopterin S-methyltransferase subunit B	891.5	1229.14	1409.83	125.58	163.91	69.24
ERBA_00729	Coenzyme transport and metabolism	-0.87	Coenzyme A biosynthesis bifunctional protein CoaBC	86.02	108.9	69.68	13.9	16.34	9.7
ERBA_01753	Coenzyme transport and metabolism	-0.79	Thiamine biosynthesis lipoprotein ApbE	163.99	196.5	249.8	20.07	26.07	30.92
ERBA_02340	Coenzyme transport and metabolism	-0.76	putative cobalamin biosynthesis protein CobD	179.3	211.02	223.41	30.04	13.4	29.45
ERBA_02337	Coenzyme transport and metabolism	-0.74	putative molybdenum cofactor guanylyltransferase	207.82	240.46	384.13	36.82	17.07	10.85
ERBA_01805	Coenzyme transport and metabolism	-0.66	Tetrahydromethanopterin S-methyltransferase subunit C	965.75	1069.87	1500.53	220.77	127.7	119.64
ERBA_00298	Coenzyme transport and metabolism	-0.64	Molybdenum cofactor biosynthesis protein B	278.54	301.67	381.5	33.55	35.93	62.97
ERBA_01895	Coenzyme transport and metabolism	-0.6	putative cobyric acid synthase	270.92	287.67	338.62	52.3	30.16	24.92
ERBA_02338	Coenzyme transport and metabolism	-0.57	Cobalamin synthase	200.26	206.6	327.84	19.26	14.24	23.52
ERBA_00869	Coenzyme transport and metabolism	-0.55	Malonyl-acyl-carrier protein O-methyltransferase	1118.76	1120.12	1065.89	192.06	18.69	119.61
ERBA_02222	**Energy production and conversion**	**-2.69**	**CoB--CoM heterodisulfide reductase 1 iron-sulfursubunit A**	**108.59**	**495.57**	**24.2**	**59.76**	**26.63**	**13.09**
ERBA_02223	**Energy production and conversion**	**-2.64**	**Nitrogenase iron protein**	**71.98**	**309.78**	**24.05**	**0.9**	**23.26**	**8.62**
ERBA_00115	**Energy production and conversion**	**-2.16**	**NADH-quinone oxidoreductase subunit M**	**456.69**	**1421.11**	**1001.57**	**61.34**	**115.49**	**122.88**
ERBA_00116	**Energy production and conversion**	**-2.14**	**NADH-quinone oxidoreductase subunit N**	**489.79**	**1510.38**	**944.81**	**118.71**	**83.52**	**114.32**
ERBA_00113	**Energy production and conversion**	**-1.88**	**NADH-quinone oxidoreductase subunit K**	**391.19**	**999.73**	**755.07**	**7.51**	**74.05**	**30.95**
ERBA_01746	**Energy production and conversion**	**-1.81**	**Electron transport complex protein RnfA**	**175.18**	**425.64**	**358.09**	**8.36**	**28.04**	**20.59**
ERBA_01748	**Energy production and conversion**	**-1.78**	**Electron transport complex protein RnfG**	**282.76**	**671.4**	**530.76**	**46.14**	**39.36**	**3.63**
ERBA_01747	**Energy production and conversion**	**-1.76**	**Electron transport complex protein RnfE**	**197.29**	**462.06**	**377.22**	**7.33**	**42.25**	**48.99**
ERBA_01615	**Energy production and conversion**	**-1.65**	**Coenzyme F420-dependent N-methenyltetrahydromethanopterin dehydrogenase**	**1580.25**	**3493.89**	**2572.97**	**547.82**	**209.25**	**268.14**
ERBA_00111	**Energy production and conversion**	**-1.55**	**Energy production NADH-quinone oxidoreductase subunit J**	**368.87**	**749.49**	**687.1**	**32.72**	**124.04**	**52.5**
ERBA_01679	**Energy production and conversion**	**-1.52**	**Pyruvate synthase subunit PorD**	**242.68**	**480.4**	**368.93**	**37.73**	**65.36**	**141.24**
ERBA_00112	**Energy production and conversion**	**-1.38**	**NADH-quinone oxidoreductase chain 6**	**541.79**	**978.8**	**850.65**	**73.88**	**124.33**	**111.77**
ERBA_01680	**Energy production and conversion**	**-1.28**	**2-oxoglutarate synthase subunit KorC**	**281.1**	**471.24**	**395.1**	**99.7**	**50.36**	**89.56**
ERBA_01677	**Energy production and conversion**	**-1.27**	**Pyruvate synthase subunit PorB**	**623.5**	**1029.37**	**859.57**	**181.41**	**78.25**	**146.55**
ERBA_01038	**Energy production and conversion**	**-1.22**	**Putative rubrerythrin**	**282.53**	**464.31**	**267.06**	**125.36**	**37.76**	**46.48**
ERBA_00675	**Energy production and conversion**	**-1.18**	**putative ferredoxin**	**2521.63**	**4052.02**	**5318.37**	**1396.23**	**186.07**	**356.56**
ERBA_00110	**Energy production and conversion**	**-1.16**	**NADH-quinone oxidoreductase subunit I**	**570.94**	**882.55**	**856.38**	**38.5**	**44.66**	**28.75**
ERBA_01749	**Energy production**	**-1.12**	**Electron transport complex**	**396.89**	**598.58**	**604.5**	**11.84**	**68.14**	**92.05**

	and conversion		**protein RnfD**						
ERBA_01678	Energy production and conversion	-0.98	Pyruvate synthase subunit PorA	543.65	736.29	697.22	148.91	27.88	160.15
ERBA_00109	Energy production and conversion	-0.97	F oxidoreductase subunit H	765.19	1037.34	1050.78	109.51	63.23	75.28
ERBA_01637	Energy production and conversion	-0.81	Tungsten-containing formylmethanofuran dehydrogenase 2 subunit C	1569	1927.62	1517.4	374.57	87.35	120.64
ERBA_02115	Energy production and conversion	-0.74	510-methylenetetrahydromethanopterin reductase	2814.91	3292.75	2655.53	573.71	152.35	253.34
ERBA_00224	Energy production and conversion	-0.62	Formylmethanofuran--tetrahydromethanopterin formyltransferase	754.32	807.39	916.47	136.75	16.97	37.84
ERBA_00108	Energy production and conversion	-0.61	NADH-quinone oxidoreductase subunit D	1053.57	1109.28	1240.91	80.43	139.55	49.89
ERBA_01750	Energy production and conversion	-0.59	Electron transport complex protein RnfC	634.62	663.29	871.68	65.72	44.94	36.99
ERBA_02044	Energy production and conversion	-0.58	tRNA-modifying enzyme	174.95	180.86	269.48	25.03	16.35	36.61
ERBA_01745	Energy production and conversion	-0.57	Electron transport complex protein RnfB	496.91	511.48	488.37	59.7	73.27	20.04
ERBA_00054	Energy production and conversion	-0.55	V-type ATP synthase subunit E	1674.83	1680.02	1236.3	299.92	206.12	69.06
ERBA_02220	**Function unknown**	**-3.75**	**dinitrogenase iron-molybdenum cofactor biosynthesi**	**44.69**	**426.2**	**7.68**	**17.71**	**72.71**	**6.65**
ERBA_02221	**Function unknown**	**-3.29**	**dinitrogenase iron-molybdenum cofactor biosynthesi**	**54.76**	**380.51**	**11.52**	**34.2**	**17.52**	**7.37**
ERBA_02217	**Function unknown**	**-2.8**	**putative ArCR protein**	**245.5**	**1186**	**195.93**	**39.86**	**41.06**	**5.43**
ERBA_00277	**Function unknown**	**-2.27**	**hypothetical protein**	**870.37**	**2993.22**	**1379.36**	**560.19**	**201.3**	**111.04**
ERBA_01160	**Function unknown**	**-2.16**	**SCP-like extracellular**	**253.45**	**793.04**	**489.7**	**73.51**	**46.47**	**49.15**
ERBA_01004	**Function unknown**	**-1.73**	**Alkyl hydroperoxide reductase AhpD**	**39.49**	**90.44**	**100.45**	**18.67**	**17.01**	**52.39**
ERBA_01492	**Function unknown**	**-1.6**	**Mitochondrial biogenesis AIM24 protein**	**140.06**	**293.25**	**185.4**	**10.15**	**21.86**	**29.82**
ERBA_00957	**Function unknown**	**-1.53**	**hypothetical protein**	**574.75**	**1166.77**	**710.98**	**207.84**	**25.9**	**45.64**
ERBA_02143	**Function unknown**	**-1.33**	**putative ACR protein**	**94.8**	**166.17**	**97.36**	**24.55**	**46.65**	**13.7**
ERBA_01413	**Function unknown**	**-1.23**	**Sirohydrochlorin cobaltochelatase**	**379.6**	**624.24**	**410.02**	**102.16**	**91.54**	**41.95**
ERBA_01466	**Function unknown**	**-1.21**	**hypothetical protein**	**106.99**	**173.54**	**174.26**	**29.49**	**25.99**	**73.32**
ERBA_01951	**Function unknown**	**-1.12**	**Iron-sulfur cluster insertion protein ErpA**	**1851.98**	**2831.08**	**2481.6**	**640.27**	**328.69**	**278.81**
ERBA_01419	**Function unknown**	**-1.11**	**Integral membrane DUF95 protein**	**82.48**	**123.76**	**108.47**	**9.88**	**12.87**	**10.82**
ERBA_02151	**Function unknown**	**-1.05**	**Protein of unknown function DUF111**	**121.94**	**175.33**	**193.06**	**7.5**	**5.2**	**9.37**
ERBA_02103	**Function unknown**	**-1.05**	**Segregation**	**160.76**	**229.29**	**150.6**	**36.78**	**29.91**	**22.66**
ERBA_01401	**Function unknown**	**-1**	**Protein of unknown function DUF116**	**182.62**	**255.19**	**383.63**	**43.53**	**26.14**	**84.28**
ERBA_01761	Function unknown	-0.94	hypothetical protein	71.41	95.08	99.91	12.74	9.51	23.74
ERBA_01418	Function unknown	-0.93	Membrane of protein	129.97	172.44	166.81	33.51	11.57	31.05
ERBA_00245	Function unknown	-0.91	methyltransferase	548.55	715.55	509.65	122.91	97.18	78.06
ERBA_01465	Function unknown	-0.88	hypothetical protein	323.52	415.04	335.63	80.64	66.76	28.73
ERBA_01400	Function unknown	-0.87	Protein of unknown function DUF116	104.45	133.4	143.17	14.37	36.36	20.97
ERBA_01416	Function unknown	-0.81	hypothetical protein	130.4	159.68	212.4	28.5	10.01	13.37
ERBA_01744	Function unknown	-0.81	Tic20-like protein	205.41	246.55	228.54	45.13	48.94	53.52
ERBA_01264	Function unknown	-0.7	hypothetical protein	272.24	307.42	197.62	42.87	34.15	48.12
ERBA_00223	Function unknown	-0.68	TraB family protein	382.44	421.73	517.9	11.66	14.72	20.23
ERBA_01168	Function unknown	-0.67	Integral membrane DUF92 protein	123.58	134.81	120.1	27.99	25.8	26.62
ERBA_01083	Function unknown	-0.66	hypothetical protein	125	136.81	155.53	28.33	19.13	13.91
ERBA_00920	Function unknown	-0.6	transmembrane protein	227.23	238.19	209.32	27.76	8.72	29.59
ERBA_02084	Function unknown	-0.51	D-tyrosyl-tRNA deacylase	289.28	284.06	368.2	15.51	21.27	18.96
ERBA_01976	Function unknown	-0.46	hypothetical protein	408.67	388.28	359.13	28.27	7.83	5.53
ERBA_00494	Function unknown /Inorganic ion transport and metabolism	-0.48	potassium channel protein	213.43	206.95	235.95	18.8	1.48	5.48
ERBA_02219	**General function prediction only**	**-3.14**	**HTH DNA-binding protein**	**165.82**	**1024.3**	**54.8**	**53.78**	**42.19**	**14.69**
ERBA_01732	**General function prediction only**	**-2.4**	**Trimethylamine corrinoid protein 2**	**4363.07**	**16109.65**	**4265.62**	**1369.17**	**1785.56**	**365.49**
ERBA_02224	**General function prediction only**	**-2.35**	**hypothetical protein**	**75.78**	**264.53**	**26.77**	**23.73**	**59.56**	**5.65**
ERBA_01728	**General function prediction only**	**-2.07**	**Dimethylamine corrinoid protein 2**	**1297.81**	**3805.85**	**1993.76**	**303.77**	**373.43**	**100.46**
ERBA_01135	**General function prediction only**	**-1.63**	**Monomethylamine corrinoid protein 1**	**2621.75**	**5776.68**	**4330.98**	**1526.11**	**483.84**	**312.6**
ERBA_01640	**General function prediction only**	**-1.5**	**S-layer-related protein**	**10198.99**	**20069.12**	**12638.49**	**1735.89**	**1715.91**	**1391.46**
ERBA_02225	**General function prediction only**	**-1.41**	**Hydroxyacylglutathione hydrolase**	**119.27**	**216.22**	**47.9**	**44.66**	**37.6**	**14.26**
ERBA_02061	**General function prediction only**	**-1.4**	**Rubredoxin-like zinc ribbon domain protein**	**148.24**	**268.39**	**142.95**	**38.95**	**28.42**	**20.38**
ERBA_01658	**General function prediction only**	**-1.4**	**Molybdenum transport protein ModE**	**387.57**	**701.8**	**639.93**	**67.92**	**129.03**	**28.49**
ERBA_02150	**General function prediction only**	**-1.29**	**Translation initiation factor 2 subunit beta**	**1708.7**	**2977.97**	**1383.2**	**1064.1**	**466.57**	**58.63**
ERBA_02149	**General function prediction only**	**-1.23**	**Translation initiation factor 2 subunit beta**	**2700.53**	**4507.03**	**2197.16**	**1572.62**	**609.37**	**336.27**
ERBA_00512	**General function prediction only**	**-1.14**	**putative GTP-binding protein EngB**	**100.51**	**152.91**	**119.06**	**11.9**	**22.46**	**14.76**

ERBA_01357	**General function prediction only**	**-1.11**	**protein of unknown function DUF20 protein**	**187.33**	**281.02**	**159.56**	**41.65**	**32.16**	**12.73**
ERBA_02154	**General function prediction only**	**-1.09**	**phosphoglycolate phosphatase**	**620.07**	**907.18**	**498.28**	**37.39**	**110.57**	**8.75**
ERBA_00261	**General function prediction only**	**-1.08**	**Rhomboid protease GlpG**	**125.63**	**185.25**	**240.98**	**23.33**	**28.48**	**17.55**
ERBA_00007	General function prediction only	-0.99	hypothetical protein	400.45	553.76	332.82	51.64	46.38	79.79
ERBA_01420	General function prediction only	-0.98	5-formaminoimidazole-4-carboxamide-1--D-ribofuranosyl 5'-monophosphate syn	161.54	220.92	280.18	21.69	6.2	32.08
ERBA_00717	General function prediction only	-0.96	Putative CRISPR-associated nuclease-helicase Cas3	41.28	55.31	55.9	2.64	3.87	14.46
ERBA_01741	General function prediction only	-0.95	CoB--CoM heterodisulfide reductase 1 iron-sulfursubunit A	984.67	1356.12	1666.82	576.61	145.12	136.46
ERBA_02136	General function prediction only	-0.84	Cyclic pyranopterin monophosphate synthase	597.12	748.26	882.45	148.47	3.55	99.08
ERBA_00921	General function prediction only	-0.83	6-hydroxymethyl-78-dihydropterin pyrophosphokinase	124.81	153.95	143.28	10.59	9.62	12.33
ERBA_00251	General function prediction only	-0.77	Carbamoyl-phosphate synthase large chain	192.12	224.22	262.77	39.12	5.07	30.99
ERBA_00864	General function prediction only	-0.73	Protein of unknown function DUF134	1558.34	1791.8	443.89	157.41	96.67	16.97
ERBA_01246	General function prediction only	-0.72	iron-sulfur flavoprotein	161.63	184.85	139.05	5.92	21.27	18.07
ERBA_00325	General function prediction only	-0.71	metal-dependent phosphohydrolase	93.84	106.48	225.31	5.81	18.07	11.47
ERBA_01251	General function prediction only	-0.67	V4R domain protein	93.1	101.89	126.43	16.93	4.6	20.66
ERBA_01302	General function prediction only	-0.65	Pyridoxine-pyridoxamine 5'-phosphate oxidase	411.33	447.62	457.19	32.24	10.23	56.62
ERBA_01467	General function prediction only	-0.65	protein of unknown function DUF87 protein	342.03	371.61	327.86	24.73	31.7	25.25
ERBA_00933	General function prediction only	-0.63	Phosphate acetyltransferase	164.04	174.75	172.98	20.4	20.52	20.6
ERBA_01182	General function prediction only	-0.52	Phosphoribosylformylglycinamidine synthase 2	289.32	285.41	397.22	22.54	23.25	30.36
ERBA_01429	General function prediction only / Amino acid transport and metabolism	-0.8	putative cobalt-precorrin-3B C-methyltransferase	149.24	180.76	167.91	26.09	22.2	42.07
ERBA_01735	General function prediction only / Energy production and conversion	-0.54	CoB--CoM heterodisulfide reductase iron-sulfur subunit A	170.58	172.51	325.35	17.21	3.26	32.42
ERBA_00222	**General function prediction only /Transcription**	**-1.09**	**tRNA nucleotidyltransferase CCA-adding enzyme**	**241.15**	**357.08**	**430.59**	**28.67**	**37.99**	**42.71**
ERBA_01661	**Inorganic ion transport and metabolism**	**-3.1**	**Molybdate-tungstate-binding protein WtpA**	**1282.54**	**7705.52**	**5819.05**	**416.08**	**327.71**	**602.59**
ERBA_00860	**Inorganic ion transport and metabolism**	**-3.01**	**Putative antiporter subunit mnhC2**	**28.25**	**158.05**	**64.37**	**10.73**	**5.63**	**11.08**
ERBA_00861	**Inorganic ion transport and metabolism**	**-2.44**	**Putative antiporter subunit mnhB2**	**48.05**	**181.47**	**111.31**	**12.6**	**37.2**	**19.72**
ERBA_00858	**Inorganic ion transport and metabolism**	**-1.83**	**Na antiporter subunit E1**	**72.82**	**178.11**	**94.17**	**8.18**	**19.91**	**14.79**
ERBA_00856	**Inorganic ion transport and metabolism**	**-1.35**	**Na antiporter subunit G1**	**52.56**	**92.29**	**64.94**	**27.5**	**4.89**	**10.2**
ERBA_01124	**Inorganic ion transport and metabolism**	**-1.03**	**putative cation-transporting ATPase F**	**98.34**	**138.92**	**45.1**	**0.53**	**11.83**	**2.62**
ERBA_00020	Inorganic ion transport and metabolism	-0.81	Phosphate import ATP-binding protein PstB	303.09	365.37	330.34	32.77	28.64	36.42
ERBA_02232	Inorganic ion transport and metabolism	-0.74	Manganese-binding lipoprotein MntA	132.72	153.76	59.92	32.95	40.09	6.55
ERBA_01979	Inorganic ion transport and metabolism	-0.61	Phosphate-specific transport system accessory protein PhoU	1125.2	1196.2	938.57	169.37	53.06	33.39
ERBA_02046	Intracellular trafficking, secretion, and vesicular transport	-0.94	peptidase S26B signal peptidase	333.19	442.8	298.45	29.82	26	27.78
ERBA_02060	**Lipid transport and metabolism**	**-1.31**	**putative acetyl-CoA acyltransferase**	**189.76**	**322.3**	**157.09**	**67.84**	**28.76**	**18.66**
ERBA_02059	**Lipid transport and metabolism**	**-1.31**	**hydroxymethylglutaryl-CoA synthase**	**183.34**	**310.93**	**148.49**	**53.95**	**8.85**	**18.82**
ERBA_02339	Lipid transport and metabolism	-0.71	alpha-ribazole phosphatase CobZ	273.91	311.9	471.26	39.88	52.48	48.79
ERBA_02126	Lipid transport and metabolism	-0.62	myo-inositol-1-phosphate synthase	374.81	400.01	607.26	44.53	21.44	48.05
ERBA_00795	**Multiple classes**	**-3.86**	**Chemotaxis protein CheW**	**134.79**	**1370.69**	**215.67**	**38.13**	**191.11**	**13.78**
ERBA_00800	**Multiple classes**	**-3.78**	**Chemotaxis protein methyltransferase**	**36.21**	**345.85**	**66.62**	**4.33**	**41.78**	**15.18**
ERBA_00798	**Multiple classes**	**-3.65**	**Chemotaxis protein CheA**	**74.83**	**648.57**	**119.51**	**19.22**	**91.04**	**15.3**
ERBA_00797	**Multiple classes**	**-3.54**	**Chemotaxis response**	**108.76**	**874.29**	**165.64**	**31.19**	**125.16**	**12.37**

			regulator protein-glutamate methylesterase						
ERBA_00794	**Multiple classes**	**-3.42**	**Methyl-accepting chemotaxis protein TlpA**	**227.82**	**1681.66**	**279.58**	**57.87**	**191.63**	**3.14**
ERBA_00799	**Multiple classes**	**-3.4**	**Flagellar motor switch phosphatase FliY**	**67.91**	**498.98**	**104.82**	**27.25**	**68.96**	**7.82**
ERBA_00801	**Multiple classes**	**-3.1**	**Flagellar motor switch phosphatase FliY**	**71.03**	**420.06**	**85.74**	**24.21**	**47.23**	**20.92**
ERBA_00332	**Multiple classes**	**-2.43**	**type II secretion system protein**	**40.34**	**149.3**	**39.87**	**8**	**14.31**	**4.95**
ERBA_00859	**Multiple classes**	**-2.12**	**Na antiporter subunit D1**	**70.84**	**210.84**	**97.69**	**18.71**	**20.82**	**6.07**
ERBA_00114	**Multiple classes**	**-1.95**	**NADH-quinone oxidoreductase subunit L**	**469.79**	**1265.41**	**1053.77**	**88.89**	**118.71**	**91.63**
ERBA_00333	**Multiple classes**	**-1.67**	**type II secretion system protein E**	**128.11**	**280.63**	**98.27**	**23.17**	**9.97**	**3.45**
ERBA_01241	**Multiple classes**	**-1.39**	**Phosphate regulon transcriptional regulatory protein PhoB**	**339.75**	**624.58**	**348.98**	**82.08**	**29.06**	**10.86**
ERBA_00862	**Multiple classes**	**-1.37**	**Na antiporter subunit A1**	**191.64**	**343.3**	**301.79**	**17.79**	**21.49**	**35.84**
ERBA_01788	**Multiple classes**	**-1.34**	**Anthranilate synthase component 1 2**	**97.32**	**172.1**	**136.19**	**21.17**	**11.24**	**18.8**
ERBA_02199	**Multiple classes**	**-1.13**	**Polar-differentiation response regulator DivK**	**289.03**	**442.48**	**297.68**	**85.56**	**14.4**	**10.51**
ERBA_01237	**Multiple classes**	**-1.08**	**Circadian clock protein kinase KaiC**	**123.61**	**179.14**	**66.25**	**32.43**	**20.92**	**12.77**
ERBA_00449	**Multiple classes**	**-1**	**Preflagellin peptidase**	**46.93**	**65.33**	**66.32**	**5.83**	**8.82**	**12.07**
ERBA_02021	Multiple classes	-0.99	Sodium-proline symporter	135.28	184.34	126.96	29.82	19.11	7.6
ERBA_01974	Multiple classes	-0.99	RIO-type serine-threonine-protein kinase Rio1	547.8	756.62	690.71	91.88	66.68	94.58
ERBA_01236	Multiple classes	-0.82	type II secretion system protein E	159.93	193.04	65.75	42.46	25.06	12.89
ERBA_01235	Multiple classes	-0.79	flagellar assembly protein J	93.4	110.26	65.94	28.39	10.51	5.06
ERBA_00454	Multiple classes	-0.6	4-hydroxy-tetrahydrodipicolinate synthase	460.34	486.95	697.98	79.02	58.61	71.12
ERBA_02188	Multiple classes	-0.59	DNA excision repair protein ERCC-2	101.47	105.53	163.66	12.98	5.94	8.42
ERBA_00554	Multiple classes	-0.53	Shikimate kinase	327.42	325.71	311.64	40.07	10.15	20.38
ERBA_02282	**Nucleotide transport and metabolism**	**-1.64**	**Deoxyguanosinetriphosphate triphosphohydrolase-like protein**	**72.92**	**159.8**	**116.66**	**29.93**	**21.29**	**15.92**
ERBA_00451	**Nucleotide transport and metabolism**	**-1.26**	**Aspartate carbamoyltransferase**	**171.87**	**285.45**	**213.12**	**14.29**	**29.84**	**11.22**
ERBA_02139	Nucleotide transport and metabolism	-0.91	Phosphoribosylaminoimidazole-succinocarboxamide synthase	244.86	319.59	154.85	35.27	22.13	3.24
ERBA_02062	Nucleotide transport and metabolism	-0.91	adenylate cyclase	215.05	276.2	145.82	72.76	34.83	21.62
ERBA_00452	Nucleotide transport and metabolism	-0.86	Aspartate carbamoyltransferase regulatory chain	154.6	192.68	157.95	30.54	9.68	36.29
ERBA_00765	Nucleotide transport and metabolism	-0.77	Adenylate kinase	647.9	757.47	743.29	106.06	56.62	27.96
ERBA_00763	Nucleotide transport and metabolism	-0.69	Cytidylate kinase	339.09	373.12	404.81	103.16	35.47	22.31
ERBA_01660	**Posttranslational modification, protein turnover, chaperones**	**-2.78**	**Molybdate-tungstate transport system permease protein WtpB**	**266.49**	**1259.88**	**1232.56**	**29.69**	**69.77**	**35.93**
ERBA_00863	Posttranslational modification, protein turnover, chaperones	-0.77	Small heat shock protein C4	1105.23	1328.22	1140.97	433.33	66.92	63.87
ERBA_02041	Posttranslational modification, protein turnover, chaperones	-0.77	FKBP-type peptidyl-prolyl cis-trans isomerase 2 Posttranslational modification	1840.55	2197.33	3240.6	403.9	136.77	185.83
ERBA_01873	Posttranslational modification, protein turnover, chaperones	-0.63	Proteasome-activating nucleotidase	969.45	1045.72	863.17	225.18	47.68	56.49
ERBA_01681	Posttranslational modification, protein turnover, chaperones	-0.48	ATP-dependent 26S proteasome regulatory subunit Posttranslational modification	184.22	177.15	160.52	2.22	14.81	2.65
ERBA_00723	**Replication, recombination and repair**	**-1.95**	**CRISPR-associated exonuclease Cas4**	**61.68**	**163.44**	**89.86**	**10.41**	**18.82**	**17.18**
ERBA_00721	**Replication, recombination and repair**	**-1.75**	**CRISPR-associated endoribonuclease Cas2**	**29.54**	**70.47**	**36.96**	**22.1**	**23.69**	**10.23**
ERBA_00724	**Replication, recombination and repair**	**-1.74**	**CRISPR-associated endoribonuclease Cas6 1**	**127.1**	**295.18**	**166.04**	**24.26**	**21.5**	**25.65**
ERBA_00722	**Replication, recombination and repair**	**-1.35**	**CRISPR-associated protein Cas1 1**	**88.14**	**154.16**	**101.41**	**24.41**	**14.69**	**12.5**
ERBA_02155	**Replication, recombination and repair**	**-1.35**	**DNA repair**	**1532.49**	**2721.12**	**1239.42**	**240.25**	**115.56**	**99.01**
ERBA_00718	**Replication, recombination and repair**	**-1.3**	**putative DNA repair protein**	**71.15**	**121.77**	**116.92**	**14.33**	**15.53**	**22.19**
ERBA_00719	**Replication, recombination and repair**	**-1.26**	**DevR family CRISPR-associated autoregulator**	**105.6**	**175.27**	**243.21**	**23.63**	**16.61**	**26.09**
ERBA_00958	**Replication, recombination and**	**-1.25**	**replication factor A**	**857.53**	**1430.62**	**944.93**	**291.24**	**55.63**	**27.6**

	repair								
ERBA_02277	Replication, recombination and repair	-0.95	DNA helicase II - ATP-dependent DNA helicase PcrA	78.11	104.17	93.21	3.01	0.35	3.22
ERBA_00331	Replication, recombination and repair	-0.94	Ribonuclease HII	83.77	110.89	71.94	12.81	32.73	6.97
ERBA_01199	Replication, recombination and repair	-0.89	DNA polymerase sliding clamp	1207.81	1566.25	1718.21	323.02	136.61	58.16
ERBA_00183	Replication, recombination and repair	-0.88	endonuclease III	127.83	161.76	179.68	39.36	22.55	10.63
ERBA_00476	Replication, recombination and repair	-0.78	Endonuclease III	233.94	281.73	447.08	59.49	45.79	30.35
ERBA_01743	Replication, recombination and repair	-0.71	Replication factor C small subunit	197.48	221.81	196.55	26.58	5.52	31.58
ERBA_01601	Replication, recombination and repair	-0.61	hypothetical protein	371.21	389.28	330.22	35.43	26.96	34.01
ERBA_00364	Replication, recombination and repair	-0.6	replicative DNA helicase Mcm	124.04	128.48	142.11	23.9	9.62	17.65
ERBA_01327	Replication, recombination and repair	-0.59	Putative transposase in snaA-snaB intergenic region	433.45	449.69	156.98	122.21	19.77	11.26
ERBA_00102	Secondary metabolites biosynthesis, transport and catabolism	-0.83	Phytoene desaturase	100.23	123.73	142.11	17.81	11.95	15.68
ERBA_00796	**Signal transduction mechanisms**	**-3.71**	**Chemotaxis response regulator protein-glutamate methylesterase**	**114.72**	**1049.52**	**162.25**	**11.05**	**143.57**	**5.3**
ERBA_00599	**Signal transduction mechanisms**	**-1.79**	**putative circadian clock protein KaiC**	**168.58**	**402.61**	**216.03**	**5.5**	**14.41**	**16.4**
ERBA_01274	**Signal transduction mechanisms**	**-1.42**	**multisensor signal transduction histidine kinase**	**110.38**	**207.51**	**162.26**	**35.69**	**13.12**	**17.14**
ERBA_02035	**Signal transduction mechanisms**	**-1.41**	**Circadian clock protein kinase KaiC**	**457.8**	**849.55**	**630.97**	**112.73**	**49.59**	**74.98**
ERBA_00918	**Signal transduction mechanisms**	**-1.32**	**KaiC domain containing protein**	**270.95**	**471.91**	**457.84**	**58.66**	**44.71**	**39.11**
ERBA_01949	**Signal transduction mechanisms**	**-1.31**	**Polar-differentiation response regulator DivK**	**222.16**	**388.05**	**96.44**	**97.36**	**46.46**	**39.23**
ERBA_01594	**Signal transduction mechanisms**	**-1.05**	**Sensor histidine kinase ResE**	**222.79**	**321.01**	**334.84**	**42.32**	**16.88**	**16.1**
ERBA_01166	Signal transduction mechanisms	-0.77	Sporulation kinase C	67.24	78.71	79.64	13.67	8	5.93
ERBA_01348	Signal transduction mechanisms	-0.71	Adaptive-response sensory-kinase SasA	92.42	105.72	158.22	25.9	8.27	8.71
ERBA_02336	Signal transduction mechanisms	-0.66	Circadian clock protein kinase KaiC	485.32	535.01	601.65	88.47	36.68	28.27
ERBA_00884	Signal transduction mechanisms	-0.55	Blue-light-activated histidine kinase	61.56	61.79	63.16	10.26	3.58	12.12
ERBA_01165	Signal transduction mechanisms	-0.55	Sensor histidine kinase ResE	124.24	125.83	149.18	8.26	14.07	19.04
ERBA_01675	Signal transduction mechanisms	-0.51	Sensor protein RcsC	102.44	100.25	121.59	16.41	3.72	18.09
ERBA_01982	Signal transduction mechanisms /Multiple classes	-0.68	Sporulation kinase C	73.68	81.38	117.23	1.2	9.79	7.1
ERBA_02058	**Transcription**	**-1.99**	**XRE family transcriptional regulator**	**163.26**	**446.8**	**138.01**	**49.56**	**70.73**	**19.49**
ERBA_02231	**Transcription**	**-1.37**	**Transcriptional regulator MntR**	**123.13**	**220.87**	**127.14**	**53.08**	**40.78**	**10.19**
ERBA_01952	**Transcription**	**-1.27**	**Nascent polypeptide-associated complex protein**	**1416.76**	**2402.72**	**1908.51**	**543.97**	**192.22**	**120.05**
ERBA_01593	**Transcription**	**-1.1**	**ArsR family transcriptional regulator**	**126.19**	**187.17**	**159.56**	**24**	**13.78**	**33.01**
ERBA_02137	**Transcription**	**-1.03**	**Cadmium resistance transcriptional regulatory protein CadC**	**120.59**	**172.54**	**214.75**	**44.52**	**20.52**	**55.96**
ERBA_02138	Transcription	-0.97	Putative nickel-responsive regulator	111.97	152.65	137.53	5.29	11.64	40.77
ERBA_00683	Transcription	-0.88	putative transcription termination protein NusA	458.65	588.67	717.41	85.32	31.61	28.1
ERBA_00685	Transcription	-0.64	DNA-directed RNA polymerase subunit A"	460.84	493.04	519.91	107.45	41.58	69.71
ERBA_00021	Transcription	-0.63	Inosine-5'-monophosphate dehydrogenase	847.59	912.98	1039.56	121.77	38.91	48.16
ERBA_01597	**Translation, ribosomal structure and biogenesis**	**-1.18**	**Translation initiation factor 1A 2**	**641.78**	**1030.88**	**1056.76**	**240.76**	**351.93**	**101.36**
ERBA_01973	**Translation, ribosomal structure and biogenesis**	**-1.08**	**Translation initiation factor 1A 1**	**424.85**	**624.41**	**641.26**	**70.15**	**90.77**	**141.81**
ERBA_00252	**Translation, ribosomal structure and biogenesis**	**-1.05**	**NMD3 family protein**	**442.35**	**634.57**	**589.48**	**33.77**	**32.47**	**44.65**
ERBA_00767	Translation, ribosomal structure and biogenesis	-0.83	50S ribosomal protein L15P	547.93	669.47	550.44	85.41	105.66	99.75

ERBA_02142	Translation, ribosomal structure and biogenesis	-0.81	Protein translation factor SUI1	822.12	1003.52	1037.89	140.23	62.68	47.45
ERBA_00678	Translation, ribosomal structure and biogenesis	-0.75	30S ribosomal protein S10	1624.47	1901.5	1553.8	322.37	169.87	103.84
ERBA_00680	Translation, ribosomal structure and biogenesis	-0.72	Elongation factor 2	1498.27	1708.45	1593.22	151.39	183.75	102.38
ERBA_00681	Translation, ribosomal structure and biogenesis	-0.67	30S ribosomal protein S7	1020.43	1128.07	1341.24	191.18	81.71	17.02
ERBA_00679	Translation, ribosomal structure and biogenesis	-0.63	Elongation factor 1-alpha	3434.37	3700.32	3496.6	628.06	439.5	247.18
ERBA_00852	Translation, ribosomal structure and biogenesis	-0.61	putative tRNA threonylcarbamoyladenosine biosynthesis protein YwlC	307.37	328.86	423.15	67.25	32.42	36.93
ERBA_00876	Translation, ribosomal structure and biogenesis	-0.61	Cysteine--tRNA ligase	83.26	87.62	199.82	5.69	5.5	36.75
ERBA_00684	Translation, ribosomal structure and biogenesis	-0.58	50S ribosomal protein L30e	708.45	736.81	1047.43	164.62	106.56	35.3
ERBA_01074	Translation, ribosomal structure and biogenesis	-0.58	50S ribosomal protein L40e	939.07	974.41	880.91	83.91	51.3	96.2
ERBA_00946	Translation, ribosomal structure and biogenesis	-0.57	Leucine--tRNA ligase	253.07	259.94	351.08	3.39	12.68	15.06
ERBA_00768	Translation, ribosomal structure and biogenesis	-0.55	50S ribosomal protein L30P	648.17	652.22	589.09	78.75	107.95	85.98
ERBA_02216	Translation, ribosomal structure and biogenesis	-0.49	Methionine aminopeptidase	231.62	224.75	267.16	11.7	8.45	3.93
ERBA_00606		**-3.08**	**hypothetical protein**	**112.01**	**664.87**	**361.47**	**43.83**	**59.03**	**62.49**
ERBA_00802		**-2.95**	**hypothetical protein**	**57.12**	**304.96**	**82.01**	**11.1**	**35.87**	**7.84**
ERBA_00276		**-2.93**	**hypothetical protein**	**510.29**	**2788.5**	**1016.58**	**389.74**	**101.54**	**255.11**
ERBA_00286		**-2.56**	**hypothetical protein**	**65.79**	**270.87**	**87.92**	**7.81**	**29.02**	**25.07**
ERBA_01556		**-2.52**	**chromosomal protein**	**938.59**	**3846.93**	**1222.96**	**563.34**	**312.47**	**67.47**
ERBA_02283		**-2.44**	**hypothetical protein**	**111.56**	**426.34**	**167.44**	**42.11**	**44.59**	**36.26**
ERBA_00598		**-2.27**	**hypothetical protein**	**59.2**	**196.81**	**92.5**	**8.6**	**6.9**	**10.01**
ERBA_00285		**-2.23**	**hypothetical protein**	**47.97**	**155.11**	**50.53**	**14.96**	**27.67**	**16.2**
ERBA_01731		**-2.23**	**trimethylamine permease**	**1573.26**	**5105.33**	**2408.13**	**65.24**	**345.85**	**437.51**
ERBA_00597		**-2.18**	**hypothetical protein**	**73.19**	**227.07**	**92.02**	**12.43**	**6.18**	**6.99**
ERBA_01227		**-2.15**	**hypothetical protein**	**162.06**	**508.32**	**297.23**	**30.07**	**108.62**	**58.1**
ERBA_01768		**-2.14**	**hypothetical protein**	**57.77**	**182.17**	**116.09**	**15.14**	**63.89**	**60.65**
ERBA_00284		**-2.13**	**hypothetical protein**	**40.96**	**125.05**	**56.22**	**11.76**	**15.21**	**18.31**
ERBA_00117		**-2.1**	**F dehydrogenase subunit FpoO**	**506.11**	**1508.24**	**832.7**	**125.58**	**135.01**	**149.83**
ERBA_01730		**-2.09**	**hypothetical protein**	**1607.86**	**4747.66**	**2108.9**	**111.43**	**473.84**	**418.8**
ERBA_01005		**-2.08**	**transcriptional regulator**	**41.23**	**119.7**	**102.98**	**27.57**	**5.04**	**12.46**
ERBA_00605		**-2.04**	**hypothetical protein**	**61.47**	**173.99**	**67**	**20.59**	**27.84**	**12.86**
ERBA_01729		**-2.04**	**hypothetical protein**	**1013.74**	**2872.94**	**1442.15**	**122.58**	**118.07**	**241.79**
ERBA_00804		**-1.97**	**hypothetical protein**	**223.74**	**613.38**	**342.85**	**69.67**	**46.85**	**21.2**
ERBA_01919		**-1.95**	**hypothetical protein**	**56.74**	**151.1**	**75.11**	**24.52**	**20.88**	**17.46**
ERBA_01918		**-1.85**	**Translocon-associated beta protein**	**141.05**	**351.86**	**236.56**	**9.26**	**9.12**	**24**
ERBA_00803		**-1.81**	**hypothetical protein**	**250**	**614.46**	**356.28**	**74.3**	**41.96**	**26.09**
ERBA_00601		**-1.81**	**hypothetical protein**	**179.05**	**436.76**	**227.85**	**25.71**	**12.04**	**44.84**
ERBA_02025		**-1.81**	**hypothetical protein**	**156.06**	**388.36**	**193.72**	**86.04**	**68.83**	**11.43**
ERBA_01962		**-1.76**	**23S rRNA**	**2.36**	**5.58**	**7.06**	**0.34**	**1.5**	**1.46**
ERBA_01090		**-1.72**	**hypothetical protein**	**72.1**	**166.72**	**87.17**	**24.51**	**15.41**	**12.89**
ERBA_00608		**-1.7**	**hypothetical protein**	**125.74**	**288.99**	**295.56**	**59.81**	**42.79**	**71.68**
ERBA_00600		**-1.69**	**hypothetical protein**	**141.03**	**315.75**	**143.98**	**17.39**	**42.34**	**11.7**
ERBA_01202		**-1.66**	**hypothetical protein**	**11.99**	**26.82**	**13.25**	**1.68**	**11.06**	**4.76**
ERBA_00283		**-1.64**	**hypothetical protein**	**38.73**	**84.06**	**44.69**	**5.74**	**19.37**	**11.06**
ERBA_00929		**-1.64**	**CHAD domain containing protein**	**336.59**	**722.94**	**523.82**	**28.79**	**54.14**	**47.34**
ERBA_00561		**-1.63**	**small heat shock protein**	**1178.21**	**2558.25**	**1230.78**	**308.12**	**155.57**	**49.26**
ERBA_01134		**-1.61**	**Monomethylamine methyltransferase MtmB2**	**733.57**	**1581.22**	**1129.24**	**322.28**	**57.71**	**69.49**
ERBA_00602		**-1.57**	**hypothetical protein**	**205.81**	**425.53**	**242.53**	**45.28**	**10.28**	**8.15**
ERBA_02289		**-1.56**	**hypothetical protein**	**150.71**	**310.04**	**242.85**	**27.79**	**11.43**	**46.29**
ERBA_02005		**-1.55**	**hypothetical protein**	**212.36**	**437.6**	**354.96**	**102.72**	**27.32**	**19.47**
ERBA_00524		**-1.54**	**tRNA-Ser**	**152.87**	**311.02**	**318**	**39.78**	**22.33**	**114.14**
ERBA_00262		**-1.52**	**hypothetical protein**	**115.85**	**231.95**	**221.41**	**25.4**	**14.96**	**13.31**
ERBA_01486		**-1.51**	**Dimethylamine methyltransferase MtbB1**	**152.38**	**304.57**	**164.48**	**60.48**	**59.77**	**25.72**
ERBA_02055		**-1.49**	**Putative adenylate cyclase 3**	**846.39**	**1652.25**	**1212.22**	**179.4**	**84.54**	**155.73**
ERBA_01368		**-1.49**	**hypothetical protein**	**27.7**	**53.73**	**73.99**	**2.53**	**2.81**	**13.86**
ERBA_01283		**-1.42**	**hypothetical protein**	**68.41**	**129.34**	**103.53**	**36.08**	**19.32**	**22.12**
ERBA_00089		**-1.41**	**hypothetical protein**	**95.85**	**177.64**	**146.11**	**25.82**	**24.85**	**3.38**
ERBA_00504		**-1.41**	**hypothetical protein**	**146.76**	**271.28**	**187.79**	**42.7**	**13.52**	**14.88**
ERBA_02290		**-1.4**	**glycosyl transferase group 1**	**38.76**	**70.41**	**42.08**	**17.47**	**19.08**	**19.52**
ERBA_01282		**-1.37**	**sulfotransferase**	**51.9**	**94.36**	**66.3**	**25.19**	**1.57**	**15.23**
ERBA_00960		**-1.37**	**Dimethylamine methyltransferase MtbB3**	**3300.05**	**5914.41**	**4102.77**	**409.34**	**381.35**	**190.92**
ERBA_00930		**-1.31**	**hypothetical protein**	**796.65**	**1372.44**	**1000.98**	**86.5**	**18.33**	**74.93**
ERBA_00550		**-1.29**	**hypothetical protein**	**169.78**	**288.65**	**190.98**	**21.09**	**31.11**	**33.7**
ERBA_00466		**-1.28**	**tRNA-Met**	**1417.19**	**2389.41**	**2406.61**	**326.11**	**589.98**	**857.04**
ERBA_01325		**-1.28**	**DDE superfamily endonuclease protein**	**235.21**	**402.5**	**282.18**	**88.2**	**105.7**	**55.35**
ERBA_00053		**-1.26**	**putative ATPase proteolipid**	**2306.78**	**3798.97**	**1829.98**	**376.48**	**121.76**	**60.71**

		chain						
ERBA_02006	**-1.23**	**hypothetical protein**	**283.83**	**462.94**	**392.21**	**49.55**	**89.21**	**27.03**
ERBA_02026	**-1.23**	**hypothetical protein**	**202.55**	**327.85**	**297.86**	**24.81**	**25.9**	**51.69**
ERBA_01991	**-1.23**	**Monomethylamine methyltransferase MtmB**	**3507.75**	**5799.51**	**4943.43**	**1577.14**	**339.02**	**399.57**
ERBA_02034	**-1.22**	**hypothetical protein**	**278.1**	**446.8**	**416.05**	**10.22**	**21.14**	**51.33**
ERBA_00511	**-1.22**	**protein of unknown function DUF11 protein**	**136.86**	**217.8**	**96.62**	**28.26**	**20.25**	**24.38**
ERBA_00956	**-1.17**	**hypothetical protein**	**312.37**	**485.21**	**412.93**	**30.14**	**65.87**	**43.08**
ERBA_02347	**-1.16**	**transposase IS4**	**41770.91**	**65261.13**	**27812.42**	**9789.21**	**8862.67**	**2390.01**
ERBA_02275	**-1.14**	**hypothetical protein**	**124.78**	**193.7**	**171.59**	**43.19**	**30.49**	**23.81**
ERBA_00448	**-1.14**	**hypothetical protein**	**172.06**	**268.51**	**104.88**	**50.56**	**72.78**	**41.26**
ERBA_00534	**-1.13**	**cell surface lipoprotein**	**482.8**	**744.13**	**507.21**	**150.87**	**68.06**	**54.79**
ERBA_00851	**-1.12**	**hypothetical protein**	**116.69**	**178.92**	**179.08**	**54**	**21.74**	**43.56**
ERBA_00182	**-1.11**	**hypothetical protein**	**167.91**	**250.73**	**327.21**	**19.53**	**17.02**	**34.67**
ERBA_01228	**-1.1**	**hypothetical protein**	**161.89**	**246.9**	**287.89**	**107.47**	**27.79**	**38.14**
ERBA_00282	**-1.08**	**hypothetical protein**	**54.74**	**79.76**	**47.86**	**7.62**	**8.42**	**12.02**
ERBA_00752	**-1.07**	**tRNA-Leu**	**430.86**	**629.5**	**691.5**	**71.43**	**276.22**	**288.17**
ERBA_01044	**-1.04**	**conserved repeat domain protein**	**211.81**	**301.59**	**224.76**	**5.16**	**27.21**	**19.36**
ERBA_01829	**-1.04**	**hypothetical protein**	**643.83**	**935.5**	**505.32**	**302.96**	**46.21**	**69.96**
ERBA_00189	**-1.03**	**hypothetical protein**	**125.77**	**179.39**	**136.24**	**49.35**	**27.26**	**46.24**
ERBA_02288	**-1.03**	**hypothetical protein**	**156.11**	**221.65**	**88.43**	**29.76**	**16.89**	**15.01**
ERBA_02278	**-1**	**PD-D/EXK nuclease superfamily protein**	**115.53**	**161.28**	**109.64**	**18.43**	**16.15**	**10.9**
ERBA_01089	**-1**	**hypothetical protein**	**28.41**	**39.59**	**49.31**	**4.81**	**1.7**	**17.51**
ERBA_00591	-0.98	Phospholipase D-nuclease domain containing protein	237.85	326.61	228.97	55.34	33.83	56.7
ERBA_01167	-0.98	hypothetical protein	72.32	99.37	100.35	8.3	16.13	15.24
ERBA_00609	-0.96	hypothetical protein	124.88	168.9	200.36	24.02	15.96	50.86
ERBA_01322	-0.91	PKD domain containing protein	23.12	30.29	37.54	4.62	4.88	2.66
ERBA_02291	-0.9	hypothetical protein	74.41	96.52	115.48	10.53	10.07	19.15
ERBA_01896	-0.89	hypothetical protein	297.25	385.74	500.38	62.07	38.97	30.36
ERBA_00731	-0.89	hypothetical protein	568.65	733.66	445.53	136.82	41.15	39.08
ERBA_00720	-0.88	CRISPR-associated protein CXXC CXXC region	68.11	86.24	119.24	13.53	12.35	28.28
ERBA_01207	-0.84	methyltransferase FkbM	174.46	218.43	137.4	49.11	8.29	30.52
ERBA_02292	-0.84	hypothetical protein	183.02	227.12	234.99	26.04	14.23	14.25
ERBA_01279	-0.84	agl cluster protein AglQ	199.76	245.78	265.62	12.85	13.42	20.59
ERBA_02105	-0.83	hypothetical protein	362.4	447.6	450.94	58.24	25.36	27.58
ERBA_00746	-0.82	hypothetical protein	238.99	293.87	307.77	49.64	31.68	45.29
ERBA_01988	-0.79	hypothetical protein	436.14	523.71	415.68	37.54	50.46	59.11
ERBA_01815	-0.79	hypothetical protein	163.19	195.46	183.51	22.35	21.17	12.63
ERBA_00992	-0.78	hypothetical protein	135.5	162.31	131.19	31.72	18.76	4.54
ERBA_00873	-0.76	Glutamyl-tRNA reductase	75.13	89.07	138.81	22.61	11.91	14.63
ERBA_01794	-0.74	S-layer-related duplication domain protein	363.43	417.34	316.88	44.17	39.2	26.52
ERBA_00705	-0.74	DNA ligase D 3'-phosphoesterase domain protein	141.12	163.93	232.98	20.15	39.62	33.27
ERBA_00510	-0.71	protein of unknown function DUF11 protein	232.98	261.06	171.88	43.76	10.23	10.16
ERBA_01897	-0.7	hypothetical protein	175.24	196.74	216.42	12.37	9.11	14.36
ERBA_00441	-0.69	hypothetical protein	104.28	116.94	94	6.38	17.75	11.55
ERBA_01950	-0.68	hypothetical protein	279.76	314.36	250.11	79.74	32.25	12.51
ERBA_00075	-0.67	hypothetical protein	129.44	142.92	142.18	14.54	14.46	11.93
ERBA_01665	-0.66	hypothetical protein	582.35	638.5	514.78	26.1	38.24	9.02
ERBA_00201	-0.56	hypothetical protein	113.08	115.23	107.79	10.06	10.8	13.67
ERBA_01638	-0.56	DNA polymerase beta subunit	570.46	583.27	480.14	76.53	22.61	35.22
ERBA_00555	-0.53	hypothetical protein	399.18	398.38	378.26	71.83	24.52	39.96
ERBA_00051	-0.52	H-transporting ATP synthase subunit H	2518.8	2504.06	1782.85	178.95	284.09	180.61

2.6 References

Aggarwal AK, Rodgers DW, Drottar M, Ptashne M, Harrison SC. (1988). Recognition of a DNA operator by the repressor of phage 434: a view at high resolution. *Science* **242**: 899–907.

Alam I, Antunes A, Kamau AA, alawi WB, Kalkatawi M, Stingl U, *et al.* (2013). INDIGO – INtegrated Data Warehouse of MIcrobial GenOmes with examples from the Red Sea extremophiles Hernandez-Lemus, E (ed). *PLoS ONE* **8**: e82210.

Allen MA, Lauro FM, Williams TJ, Burg D, Siddiqui KS, De Francisci D, *et al.* (2009). The genome sequence of the psychrophilic archaeon, *Methanococcoides burtonii*: the role of genome evolution in cold adaptation. *ISME J* **3**: 1012–1035.

Antunes A, Ngugi DK, Stingl U. (2011). Microbiology of the Red Sea (and other) deep-sea anoxic brine lakes. *Environmental Microbiology Reports* **3**: 416–433.

Balderston WL, Payne WJ. (1976). Inhibition of methanogenesis in salt marsh sediments and whole-cell suspensions of methanogenic bacteria by nitrogen oxides. *Appl Environ Microbiol*

32: 264–269.

Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, *et al.* (2012). SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol* **19**: 455–477.

Belay N, Sparling R, Daniels L. (1984). Dinitrogen fixation by a thermophilic methanogenic bacterium. *Nature* **312**: 286–288.

Black SL, Dawson A, Ward FB, Allen RJ. (2013). Genes required for growth at high hydrostatic pressure in *Escherichia coli* K-12 identified by genome-wide screening. Pöggeler, S (ed). *PLoS ONE* **8**: e73995.

Blom J, Albaum SP, Doppmeier D, Pühler A, Vorhölter F-J, Zakrzewski M, *et al.* (2009). EDGAR: a software framework for the comparative analysis of prokaryotic genomes. *BMC Bioinf* **10**: 154.

Bolger AM, Lohse M, Usadel B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**: 2114–2120.

Borrel G, Parisot N, Harris HMB, Peyretaillade E, Gaci N, Tottey W, *et al.* (2014). Comparative genomics highlights the unique biology of Methanomassiliicoccales, a Thermoplasmatales-related seventh order of methanogenic archaea that encodes pyrrolysine. *BMC Genomics* **15**: 679.

Campanaro S, Williams TJ, Burg DW, De Francisci D, Treu L, Lauro FM, *et al.* (2011). Temperature-dependent global gene expression in the Antarctic archaeon *Methanococcoides burtonii*. *Environ Microbiol* **13**: 2018–2038.

Darling AE, Mau B, Perna NT. (2009). Progressive Mauve: Multiple alignment of genomes with gene flux and rearrangement. *arXiv*.

de Boer PA, Crossley RE, Hand AR, Rothfield LI. (1991). The MinD protein is a membrane ATPase required for the correct placement of the Escherichia coli division site. *EMBO J* **10**: 4371–4380.

Ferry JG. (1994). Methanogenesis: ecology, physiology, biochemistry and genetics. Chapman&Hall.

Galagan JE, Nusbaum C, Roy A, Endrizzi MG, Macdonald P, FitzHugh W, *et al.* (2002). The genome of *M. acetivorans* reveals extensive metabolic and physiological diversity. *Genome Res* **12**: 532–542.

Gaston MA, Jiang R, Krzycki JA. (2011). Functional context, biosynthesis, and genetic encoding of pyrrolysine. *Current Opinion in Microbiology* **14**: 342–349.

Gaston MA, Zhang L, Green-Church KB, Krzycki JA. (2011). The complete biosynthesis of the genetically encoded amino acid pyrrolysine from lysine. *Nature* **471**: 647–650.

Goris J, Konstantinidis KT, Klappenbach JA, Coenye T, Vandamme P, Tiedje JM. (2007). DNA-DNA hybridization values and their relationship to whole-genome sequence similarities. *Int J Syst Evol Microbiol* **57**: 81–91.

Graham DE, Taylor SM, Wolf RZ, Namboori SC. (2009). Convergent evolution of coenzyme M biosynthesis in the Methanosarcinales: cysteate synthase evolved from an ancestral threonine synthase. *Biochem J* **424**: 467–478.

Graham JE, Wilkinson BJ. (1992). Staphylococcus aureus osmoregulation: roles for choline, glycine betaine, proline, and taurine. **174**: 2711–2716.

Grissa I, Vergnaud G, Pourcel C. (2008). CRISPRcompar: a website to compare clustered regularly interspaced short palindromic repeats. *Nucleic Acids Res* **36**: W145–W148.

Grissa I, Vergnaud G, Pourcel C. (2007). CRISPRFinder: a web tool to identify clustered regularly interspaced short palindromic repeats. *Nucleic Acids Res* **35**: W52–W57.

Guan Y, Hikmawan T, Antunes A, Ngugi D, Stingl U. (2015). Diversity of methanogens and sulfate-reducing bacteria in the interfaces of five deep-sea anoxic brines of the Red Sea. *Research in Microbiology*. doi:10.1016/j.resmic.2015.07.002.

Hoehler T, Gunsalus RP, McInerney MJ. (2010). Environmental constraints that limit methanogenesis. In: *Methanogens and methanogenesis in hypersaline environments*, Springer Berlin Heidelberg: Berlin, Heidelberg, pp 635–654.

Jarrell KF, Albers S-V. (2012). The archaellum: an old motility structure with a new name. *Trends in Microbiology* **20**: 307–312.

Johnson EF, Mukhopadhyay B. (2005). A new type of sulfite reductase, a novel coenzyme F_{420}-dependent enzyme, from the methanarchaeon *Methanocaldococcus jannaschii. J Biol Chem* **280**: 38776–38786.

Johnson EF, Mukhopadhyay B. (2008). Coenzyme F_{420}-dependent sulfite reductase-enabled sulfite detoxification and use of sulfite as a sole sulfur source by *Methanococcus maripaludis*. *Appl Environ Microbiol* **74**: 3591–3595.

Kerepesi C, Bánky D, Grolmusz V. (2014). AmphoraNet: the webserver implementation of the AMPHORA2 metagenomic workflow suite. *Gene* **533**: 538–540.

Kimura Y, Ohtani M, Takegawa K. (2005). An adenylyl cyclase, CyaB, acts as an osmosensor in *Myxococcus xanthus*. **187**: 3593–3598.

Kloda A, Martinac B. (2001). Structural and functional differences between two homologous mechanosensitive channels of *Methanococcus jannaschii*. *EMBO J* **20**: 1888–1896.

Koga Y, Nakano M. (2008). A dendrogram of archaea based on lipid component parts composition and its relationship to rRNA phylogeny. *Syst Appl Microbiol* **31**: 169–182.

Kopylova E, Noé L, Touzet H. (2012). SortMeRNA: fast and accurate filtering of ribosomal RNAs in metatranscriptomic data. *Bioinformatics* **28**: 3211–3217.

Lai MC, Sowers KR, ROBERTSON DE, ROBERTS MF, Gunsalus RP. (1991). Distribution of compatible solutes in the halophilic methanogenic archaebacteria. **173**: 5352–5358.

Lai S-J, Lai M-C. (2011). Characterization and regulation of the osmolyte betaine synthesizing enzymes GSMT and SDMT from halophilic methanogen Methanohalophilus portucalensis. Driessen, A (ed). *PLoS ONE* **6**: e25090.

Laine B, Culard F, Maurizot JC, Sautière P. (1991). The chromosomal protein MC1 from the archaebacterium *Methanosarcina* sp. CHTI 55 induces DNA bending and supercoiling. *Nucleic Acids Res* **19**: 3041–3045.

Lampe DJ, Churchill ME, Robertson HM. (1996). A purified mariner transposase is sufficient to mediate transposition in vitro. *EMBO J* **15**: 5470–5479.

Lange SJ, Alkhnbashi OS, Rose D, Will S, Backofen R. (2013). CRISPRmap: an automated classification of repeat conservation in prokaryotic adaptive immune systems. *Nucleic Acids Res* **41**: 8034–8044.

Langmead B, Salzberg SL. (2012). Fast gapped-read alignment with Bowtie 2. *Nat Meth* **9**: 357–359.

Levina N, Tötemeyer S, Stokes NR, Louis P, Jones MA, Booth IR. (1999). Protection of *Escherichia coli* cells against extreme turgor by activation of MscS and MscL mechanosensitive channels: identification of genes required for MscS activity. *EMBO J* **18**: 1730–1737.

Luo H, Zhang C-T, Gao F. (2014). Ori-Finder 2, an integrated tool to predict replication origins in the archaeal genomes. *Front Microbiol* **5**: 482.

Makarova KS, Haft DH, Barrangou R, Brouns SJJ, Charpentier E, Horvath P, *et al.* (2011). Evolution and classification of the CRISPR–Cas systems. *Nature Reviews Microbiology* **9**: 467–477.

Martinac B. (2004). Mechanosensitive ion channels: molecules of mechanotransduction. *J Cell Sci* **117**: 2449–2460.

Mathrani IM, Boone DR, Mah RA, Fox GE, Lau PP. (1988). *Methanohalophilus zhilinae* sp. nov., an alkaliphilic, halophilic, methylotrophic methanogen. *Int J Syst Bacteriol* **38**: 139–142.

McGenity TJ. (2010). Methanogens and Methanogenesis in Hypersaline Environments. In: *Methanogens and methanogenesis in hypersaline environments*, Springer Berlin Heidelberg: Berlin, Heidelberg, pp 665–680.

Montaño SP, Rice PA. (2011). Moving DNA around: DNA transposition and retroviral integration. *Curr Opin Struct Biol* **21**: 370–378.

Ngugi DK, Blom J, Alam I, Rashid M, Ba-Alawi W, Zhang G, *et al.* (2014). Comparative genomics reveals adaptations of a halotolerant thaumarchaeon in the interfaces of brine pools in the Red Sea. *ISME J*. doi:10.1038/ismej.2014.137.

Nurk S, Bankevich A, Antipov D, Gurevich AA, Korobeynikov A, Lapidus A, *et al.* (2013). Assembling single-cell genomes and mini-metagenomes from chimeric MDA products. *http://dxdoiorg/101089/cmb20130084* **20**: 714–737.

Nyyssola A, Kerovuo J, Kaukinen P, Weymarn von N, Reinikainen T. (2000). Extreme halophiles synthesize betaine from glycine by methylation. *J Biol Chem* **275**: 22196–22201.

Oren A. (1999). Bioenergetic aspects of halophilism. *Microbiol Mol Biol Rev* **63**: 334–348.

Oren A. (2011). Thermodynamic limits to microbial life at high salt concentrations. *Environ Microbiol* **13**: 1908–1923.

Rasko DA, Myers GSA, Ravel J. (2005). Visualization of comparative genomic analyses by BLAST score ratio. *BMC Bioinf* **6**: 2.

Richter M, Rosselló-Mora R. (2009). Shifting the genomic gold standard for the prokaryotic species definition. *Proc Natl Acad Sci USA* **106**: 19126–19131.

Robinson MD, McCarthy DJ, Smyth GK. (2010). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**: 139–140.

Ruch S, Beyer P, Ernst H, Al-Babili S. (2005). Retinal biosynthesis in Eubacteria: in vitro characterization of a novel carotenoid oxygenase from Synechocystis sp. PCC 6803. *Molecular Microbiology* **55**: 1015–1024.

Sowers KR, Baron SF, Ferry JG. (1984). *Methanosarcina acetivorans* sp. nov., an acetotrophic methane-producing bacterium isolated from marine sediments. *Appl Environ Microbiol* **47**: 971–978.

Sowers KR, Ferry JG. (1983). Isolation and characterization of a methylotrophic marine methanogen, *Methanococcoides methylutens* gen. nov., sp. nov. *Appl Environ Microbiol* **45**: 684–690.

Spring S, Scheuner C, Lapidus A, Lucas S, Glavina Del Rio T, Tice H, *et al.* (2010). The genome sequence of *Methanohalophilus mahii* SLP(T) reveals differences in the energy metabolism among members of the *Methanosarcinaceae* inhabiting freshwater and saline environments. *Archaea* **2010**: 690737.

Spudich JL, Yang CS, Jung KH, Spudich EN. (2000). Retinylidene proteins: structures and functions from archaea to humans. *Annu Rev Cell Dev Biol* **16**: 365–392.

Stamatakis A. (2006). RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* **22**: 2688–2690.

Susanti D, Wong JH, Vensel WH, Loganathan U, DeSantis R, Schmitz RA, *et al.* (2014). Thioredoxin targets fundamental processes in a methane-producing archaeon, *Methanocaldococcus jannaschii*. *Proceedings of the National Academy of Sciences* **111**: 2608–2613.

Taylor CD, Wolfe RS. (1974). Structure and Methylation of Coenzyme M ($HSCH_2CH_2SO_3$). *J Biol Chem* **249**: 4879–4885.

Thauer RK, Kaster A-K, Seedorf H, Buckel W, Hedderich R. (2008). Methanogenic archaea: ecologically relevant differences in energy conservation. *Nature Reviews Microbiology* **6**: 579–591.

Villanueva L, Damsté JSS, Schouten S. (2014). A re-evaluation of the archaeal membrane lipid biosynthetic pathway. *Nature Reviews Microbiology* **12**: 438–448.

Walker CB, la Torre de JR, Klotz MG, Urakawa H, Pinel N, Arp DJ, *et al.* (2010). *Nitrosopumilus maritimus* genome reveals unique mechanisms for nitrification and autotrophy in globally distributed marine crenarchaea. *Proceedings of the National Academy of Sciences* **107**: 8818–8823.

Wang M, Tomb J-F, Ferry JG. (2011). Electron transport in acetate-grown Methanosarcina acetivorans. *BMC Microbiol* **11**: 165.

Wu M, Scott AJ. (2012). Phylogenomic analysis of bacterial and archaeal sequences with AMPHORA2. *Bioinformatics* **28**: 1033–1034.

Wu S, Zhu Z, Fu L, Niu B, Li W. (2011). WebMGA: a customizable web server for fast metagenomic sequence analysis. *BMC Genomics* **12**: 444.

Xu S, Zhou J, Liu L, Chen J. (2011). Arginine: A novel compatible solute to protect Candida glabrata against hyperosmotic stress. *Process Biochemistry* **46**: 1230–1235.

Yelton AP, Thomas BC, Simmons SL, Wilmes P, Zemla A, Thelen MP, *et al.* (2011). A semi-quantitative, synteny-based method to improve functional predictions for hypothetical and poorly annotated bacterial and archaeal genes. Eisen, JA (ed). *PLoS Comp Biol* **7**: e1002230.

Zhilina TN, Zavarzin GA. (1987). *Methanohalobium evestigatum*, n. gen., n. sp. the extremely halophilic methanogenic archaebacterium. *Doklady Akademii nauk SSSR* **293**: 464–468.

Zhou Y, Liang Y, Lynch KH, Dennis JJ, Wishart DS. (2011). PHAST: a fast phage search tool. *Nucleic Acids Res* **39**: W347–52.

3 Chapter III CULTIVATION OF A NOVEL METHANOHALOPHILUS SPECIES FROM THE KEBRIT DEEP AND GENOMIC FEATURES OF METHANOHALOPHILUS

Yue Guan[a], David K. Ngugi[a], Manikandan Vinu[a], Sylvain Guillot[a], Jochen Blom[b], Intikhab Alam[c], James G. Ferry[d], Ulrich Stingl[a]

[a]Red Sea Research Center, King Abdullah University of Science and Technology, Thuwal, Saudi Arabia

[b]Justus-Liebig-Universität Giessen, Bioinformatik und Systembiologie, Giessen, Germany

[c]Computational Bioscience Research Center, King Abdullah University of Science and Technology, Thuwal, Saudi Arabia

[d]Department of Biochemistry and Molecular Biology, Pennsylvania State University, University Park, Pennsylvania, USA

Author contributions:

Yue Guan isolated and characterized the new species *Methanohalophilus* strain RSK, analyzed the data, and wrote this chapter. Yue Guan, David Ngugi, Manikandan Vinu performed bioinformatic analyses. Sylvain Guillot assisted with the characterization. Jochen Blom and Intikhab Alam developed and maintained two bioinformatic platforms and a genome annotation pipeline used for Chapter III. Ulrich Stingl conceived the study, which is to be submitted as an article to a peer-reviewed journal. James G. Ferry and Ulrich Stingl provided directions and lab support for the study.

Abstract

Halophilic methanogens are critical for the carbon cycle in hypersaline environments. *Methanohalophilus*-related species are ubiquitous and may contribute significantly to methane formation in the deep-sea hypersaline habitats. However, understanding the reasons for their success and their role in deep-sea hypersaline habitats has been hindered by the lack of cultures and genomes. In this chapter, I describe how a novel *Methanohalophilus* strain was isolated from the sulfide rich brine-seawater interface of Kebrit Deep in the Red Sea by using anaerobic cultivation methods. On the basis of physiological, phylogenetic and genomic features, strain RSK is considered to represent a novel species of this genus of methanogenic archaea and is the second methanogen isolated from the deep-sea hypersaline lakes of the Red Sea. We generated the draft genome sequence of the strain RSK together with *Methanohalophilus halophilus*, *Methanohalophilus euhalobius* and *Methanohalophilus portucalensis.* Our analysis revealed a highly conserved and syntenic common core genome of this genus. *De novo* synthesis of compatible solutes is conserved and crucial for halophilic methanogens to thrive in hypersaline environments. These closely related *Methanohalophilus* genomes have furthered our understanding of genomic characteristics and differentiations of halophilic methanogens.

3.1 Introduction

Methane is an important component in the global carbon cycle (Thauer *et al.*, 2008). In hypersaline environments, methanogenic processes are commonly associated with the genus of *Methanohalophilus* (Spring *et al.*, 2010). This genus currently comprises five species: *Mhp. halophilus* (Zhilina, 1983), *Mhp. euhalobius* (Bezrukova and Belyaev, 1987), *Mhp. Mahii* (Paterek and Smith, 1988), *Mhp. portucalensis* (Boone *et al.*, 1993), and *Mhp. levihalophilus* (Katayama *et al.*, 2014). They were isolated from geographically separated saline environments such as Shark Bay (Australia), saline subsurface water of oil deposits (Russia), the Great Salt Lake (USA), and a salt pond in Portugal. The *Methanohalophilus* methanogens are unique with respect to their ability to grow in a wide range of salinities (1.5% to 20% NaCl). All previously known members of this genus are mesophilic, neutrophilic, obligately methylotrophic methanogens, except for the recently described *Mhp. levihalophilus* (isolated from paleo-seawater in Japan), which appears to be a slightly halophilic methanogen. Only one genome belonging to this genus is currently publicly available (Spring *et al.*, 2010).

Deep hypersaline anoxic basins (DHABs (van der Wielen, 2005)) are some of the most remote, challenging and extreme environments on the planet. They only occur in the Mediterranean Sea, the Gulf of Mexico and the Red Sea (Antunes *et al.*, 2011). Maker gene sequences that are closely related to *Methanohalophilus* have been commonly found in deep-sea hypersaline environments (La Cono *et al.*, 2011; Yakimov *et al.*, 2013; Guan *et al.*, 2015). However, to date only one strain has been reported from Lake Thetis, eastern

Mediterranean Sea (Cono *et al.*, 2015). Due to lack of cultures, our knowledge of how *Methanohalophilus* could adapt to deep-sea hypersaline environments in the Red Sea is very limited.

The interfaces between the brines and the overlaying seawater are characterized by steep physiochemical gradients. In Kebrit Deep, within only 3m of depth difference, the salinity increases from 4% (average seawater salinity) to 26%, while the pH drops from 8.1 (seawater pH) to 5.5. Similarly, the O_2 concentration decreases from 3.2 ml/L (seawater) to zero (Eder *et al.*, 2001). The variation in chemical properties across these steep gradients provide a good model for investigating the effects of the increasingly harsh conditions on the diversity, function, and ecology of native deep-sea microbes.

While the previous chapter of this thesis describes the first strain to be ever obtained from one of the Red Sea DHABs, we describe a second strain in this chapter, a halophilic and methanogenic strain affiliated with the genus *Methanohalophilus* from the brine-sea water interface of Kebrit Deep. In this chapter, we describe the characteristics of this strain and undertake a genomic comparison among several closely related halophilic methanogens, with the aim to provide more insights into their life styles and to reveal the adaptive strategies of *Methanohalophilus* species to harsh environments.

3.2 Materials and methods

3.2.1 Sample collection

The samples from the brine-seawater interface and brines of the Kebrit Deep (24° 43' 28"N, 36° 16' 36" E) were obtained by a Rosette Sampler at depths of

1467m and 1490m below the sea surface during the R/V Aegaeo Red Sea Expedition in November 2011.

3.2.2 Enrichment and isolation

The High Salt minimal medium based on a bicarbonate buffer system previously used to isolate marine methanogens (Sowers and Ferry, 1983; Sowers *et al.*, 1984) was adapted. We increased the salinity gradient in order to target halophilic methanogens. The final pH values of the medium were adjusted from 5.5 to 7.5 with 0.5 increments. Trimethylamine (TMA), acetate, formate, or H_2:CO_2 (4:1) were supplied separately. Enrichments were initiated by adding 5ml of Kebrit Deep samples to anaerobic serum vials that contained 45ml of enrichment medium. The cultures were incubated in the dark at either 22°C, 30°C, or 37°C. Each condition was performed in triplicate. Further isolation (serial dilution, inoculation of agar shake, and plating on agar) was undertaken on cultures, if methane production was detected, and culture turbidity was increased. This process took between 1 and 3 months.

After obtaining the isolate, the effect of NaCl on growth was tested by an incremental increase of 1% NaCl in the same minimal medium, providing a range of 0% to 30% NaCl in 1% graduations. Growth characteristics were measured across a broad temperature range (4°C, 15°C, 22°C, 30°C, 33°C, 35°C, 37°C, 40°C, 45°C, and 50°C). The pH range for growth was examined from 4.0 to 9.0. H_2:CO_2 (4:1), formate, acetate, methanol, mono-, di-, and tri-methylamine, dimethylsulfide, and betaine were tested as potential growth substrates. To investigate possible growth factors and stimulatory compounds, vitamins (biotin, folic acid, pyridoxine-HCl, thimine-HCl, riboflavin, nicotinic

acid, d-Ca-pantothenate, vitamin B12, p-aminobenzoic acid, and lipoic acid), yeast extract, peptone, and trypticase were tested. All the tests were performed in triplicate.

3.2.3 Determining sensitivity to antibiotics

Ampicillin, kanamycin, carbenicillin, penicillin G, chloramphenicol, rifampicin and vancomycin were tested at a final concentration of 100 µg/ml on cultures supplied with TMA as a substrate.

3.2.4 Analytical method

Cell growth was measured as optical densities at 600nm (SpectraMax). Gas samples were removed from the incubation bottles with pressure lock syringes and were analyzed to determine the presence of methane with an Agilent gas chromatography/mass spectrometry (GC/MS).

3.2.5 Microscopy

For imaging using Transmission Electron Microscopy (TEM), cells in the exponential growth phase were fixed by adding 2.5% glutaraldehyde directly to the growth medium. Transmission electron micrographs were obtained in the Imaging and Characterization Core Lab at King Abdullah University of Science and Technology (KAUST) using a Titan G2 80-300kV TEM manufactured bf FEI. The TEM is equipped with a 4k × 4k CCD camera (US4000) and an energy filter, model GIF Tridiem manufactured by Gatan, Inc. Cyro-scanning electron microscopy (cryo-SEM) experiments were performed using a PP2000T cryo-transfer system manufactured by Qurorum Technologies, attached to an FEI Nova Nano630 SEM with a field emission electron source and through-the-lens electron detectors.

3.2.6 Whole-genome sequencing, assembly and annotation

Genomic DNA of *Methanohalophilus euhalobius* (DSM 10369), *Methanohalophilus halophilus* (DSM 3094) and *Methanohalophilus portucalensis* (DSM 7471) was provided by DSMZ (German Collection of Microorganisms and Cell Cultures). Genomic DNA of strain RSK was extracted and purified using the DNeasy Blood & Tissue Kit manufactured by Qiagen, following kit protocols.

The sequencing library was prepared using the Illumina TruSeq™ DNA Sample Preparation kit. Sequencing was done using an Illumina MiSeq sequencer at KAUST following Illumina's protocols. Reads were quality filtered, trimmed, and assembled into contigs using the *de novo* assembler SPAdes, version 2.5.1 (Bankevich *et al.*, 2012). Putative coding sequences (CDSs) of all draft genomes were predicted using the automated annotation pipeline INDIGO (www.cbrc.kaust.edu.sa/indigo; (Alam *et al.*, 2013)).

3.2.7 Comparative genomics

Overall, approaches similar to those described in Chapter II were used in order to compare and analyze the genomes of this chapter. The average nucleotide identity (ANI) was analyzed in the JSpecies software (Richter and Rosselló-Mora, 2009) and the overall gene order conservation (synteny) was evaluated by the a published method described by Yelton et al. (Yelton *et al.*, 2011). The complete genome of *Methanohalophilus mahii* (Spring *et al.*, 2010) was used as a reference for the comparative analysis. To synchronize the genome annotation, this published genome was re-annotated using the INDIGO pipeline (Alam *et al.*, 2013). Genomes were compared and evaluated on the

EDGAR platform in order to obtain information on similarities and differentiations (Blom *et al.*, 2009)

3.3 Results and discussion

3.3.1 Isolation and physiological characteristics of strain RSK

Strain RSK was isolated from an enrichment culture inoculated from the brine-seawater interface of Kebrit Deep at a depth around 1468 meters below the sea surface, where remarkably high methane, H_2S, and CO_2 content was observed (Guan *et al.*, 2015). The predominant organisms in this enrichment were cocci that occurred singly or in pairs and fluoresced blue-green when examined by UV fluorescence microscopy. Colonies on agar were white and 0.5mm in diameter after 20 days of incubation. Surface colonies were smooth, circular, and convex with entire edges. *Halanaerobium* related organisms were commonly found to appear in methanogen colonies identified by 16s RNA amplicons. These sequences were closely related to *Halanaerobium* previously isolated from the same location (Eder *et al.*, 2001). Several axenic methanogen colonies were isolated with almost identical 16S rRNA sequences. From these colonies, strain RSK was selected for further study.

Strain RSK unitizes mono-, di-, trimethylamine, and methanol as energy sources. Cell growth was not observed when the substrate was changed to H_2:CO_2 (4:1), formate, acetate or dimethyl sulfide. Strain RSK exhibited the widest salinity adaptation range for growth (2% – 20% NaCl) among all described *Methanohalophilus* species. Table 13 provides a comparison of characteristics of strain RSK and previously described species. The salinity

(NaCl) range for methane production was 2% – 25%. The temperature range for growth was 15°C – 40°C. No growth was observed at 45°C for 6 months. The most rapid growth was observed in a medium containing 5% – 10% NaCl, near 33°C and around pH 6.5. Vitamins, yeast extract or trypticase were not required and did not stimulate growth. Strain RSK was resistant to a broad range of antibiotics. Other than chloramphenicol and rifampicin, none of the other 5 commonly tested antibiotics inhibited the growth of this strain.

Table 13. Characteristics of strain RSK and other species in the genus *Methanohalophilus*.

Characteristic	*Methanohalophilus* strain RSK	*Methanohalophilus mahii*	*Methanohalophilus halophilus*	*Methanohalophilus portucalensis*	*Methanohalophilus euhalobius*
Isolation publication	This study	Paterek (1988)	Zhilina (1984)	Boone (1993)	Obraztsova (1987)
Cell type	cocci	irregular cocci	irregular cocci	irregular cocci	irregular cocci
Isolation Source	Kebrit BSI, Red Sea	Great Salt Lake, Utah, sediment	Shark Bay, Australia	Salinarium in Figiera da Foz, Portugal	Saline subsurface water, Russia, Bonduzhkoe oil deposit
Size (μm)	0.3 - 2.5	0.8 - 1.8		0.6 - 2	
Motility	non-motile	non-motile	non-motile	non-motile	non-motile
Temp. range (°C)	15-40 (not 4 or 45)	< 45	18 - 45	25-50	15-50
Optimal temp. (°C)	30-33	35 - 37	26-36	37-42	28-37
Source Temp.	21.9				
pH range	4-7.5	6.5-8.5	6.5-7.4	6.0-8.5	5.8-8.0
Optimal pH	6.5	7.5	7.4	7.2	6.8-7.3
NaCl range (%)	2 - 20	3 - 20	0.5-15	3 - 18	1-14
Optimal NaCl (%)	5 -10	12	7-9	3-13	6
Source salinity	18%				
Substrates					
TMA	+	+	+	+	+
DMA	+	+	+	+	+
MMA	+	+	+	+	+
Methanol	+	+	+	+	+
DMS	-				
Glycine betaine	-				

Acetate	-	-	-	-	-
Formate	-	-	-	-	-
H_2 CO_2	-	-	-	-	-
Growth factors	not required	not reported	not required	biotin	biotin

3.3.2 Cell morphology

Compared with published images of previously described *Methanohalophilus* species that exhibited irregular spherical shapes (Spring *et al.*, 2010), strain RSK had several distinct cellular features. The trimethylamine-grown cells were non-motile, mostly regular cocci ranging from 0.3µm to 2.5µm in diameter. The cells occurred singly, in pairs, or as small aggregates. Small nubs on the cell surface were observed, which could be extracellular vesicles or buddings. The intracellular content of strain RSK appeared to be homogeneous. No internal vesicle was observed in the thin section electron microphotograph as depicted in Figure 19.

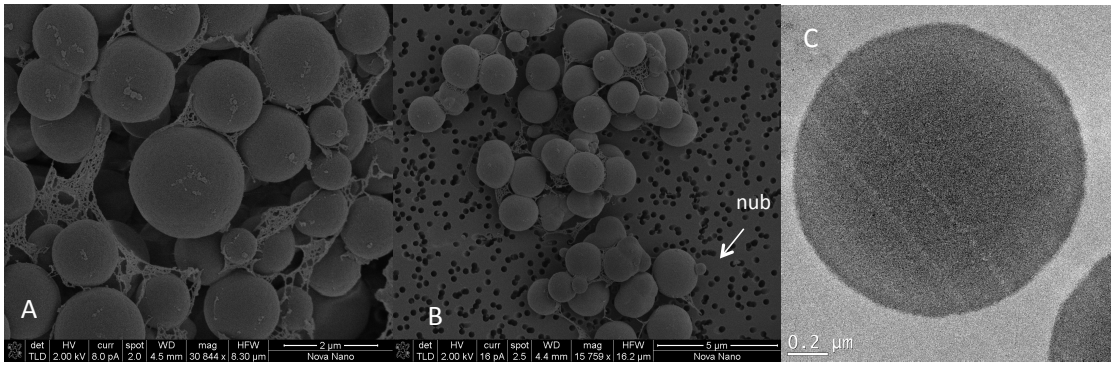


Figure 19. Electron micrograph of cells of strain RSK. Images A and B are scanning electron micrographs. Image C is a thin section electron microphotograph.

3.3.3 Comparative genomics of Methanohalophilus

An overview of genomic characteristics exhibited by members of genus *Methanohalophilus* is collected in Table 14. Based on one complete genome and four high quality draft genome assemblies of the same genus, the genome

length of *Methanohalophilus* is around 2.0Mbp, making them the smallest genomes in the family *Methanosarcinaceae*. The G+C contents of the genomes of this genus were almost identical (all around 42%). The numbers of predicted open reading frames of *Methanohalophilus* species are between 1990 and 2124. All tRNAs for 20 common proteinogenic amino acids and pyrrolysine are present.

Table 14. The assembly statistics and genome features of sequenced *Methanohalophilus* species.

Species	*Mhp.mahii*	*Mhp. halophilus*	*Mhp. portucalensis*	*Mhp. eubalobius*	Strain RSK
Reference	Spring et al, 2010	This study	This study	This study	This study
Size (bp)	2012424	2021331	2080338	1875790	1969036
No. of contigs	1	6	17	25	18
ORFs	2016	2047	2124	1990	2053
DNA G+C content (%)	42.6	42.4	41.9	42.4	41.87
rRNA operon (16S-23S-5S)	3	1	1	1	1
Unique proteins within genus Methanohalophilus (percentage)	131 (6.50%)	151 (7.38%)	216 (10.17%)	276 (13.39%)	260 (12.7%)
CRISPR repeats	0	0	1	2	3

These members of genus *Methanohalophilus* have almost identical 16S rRNA genes (>99.4% identities) and tetranucleotide signatures (>0.99). They also share very similar average nucleotide identities (ANI) as listed in Table 15. Based on an ANI threshold of less than 95%, which corresponds to a 70% DNA-DNA hybridization for identifying novel species (Goris *et al.*, 2007), a comparison among all the genomes listed in Table 3 revealed that strain RSK represents a novel species in the genus *Methanohalophilus*.

Table 15. The average nucleotide identity among genomes calculated with the BLAST algorithm.

	***Mhp.* sp. RSK**	***Mhp. portucalensis***	***Mhp. euhalobius***	***Mhp. halophilus***	***Mhp. mahii***	***Mcc. methylutens***	***Mcc. burtonii***	***Mhb. evestigatum***
***Mhp.* RSK**	---	93.57	93.04	92.32	91.08	67.75	67.19	65.77
Mhp. portucalensis	93.67	---	93.92	92.93	91.23	67.82	67.45	66.1
Mhp. euhalobius	93.27	94.1	---	92.73	91.33	67.96	67.56	66.09
Mhp. halophilus	92.51	92.99	92.72	---	91.45	67.79	67.15	65.98
Mhp. mahii	91.08	91.2	91.23	91.43	---	67.98	67.26	65.96
Mcc. methylutens	67.75	67.8	67.97	67.72	67.9	---	76.17	66.28
Mcc. burtonii	67.39	67.62	67.67	67.34	67.45	76.48	---	66.28
Mhb. evestigatum	65.9	66.01	66.06	66.12	66.17	66.35	66.12	---

Figure 20 shows a comparison of these genomes that also reveals a highly syntenic (greater than 1559 genes located in syntenic regions when compared with the completed genome of *Mhp. mahii*) common core genome, which embraces 1490 conserved CDS. Of each genome, only 6.5% to 13.4% of predicted proteins are species-specific (refer also to both Table 2 and Figure 2).

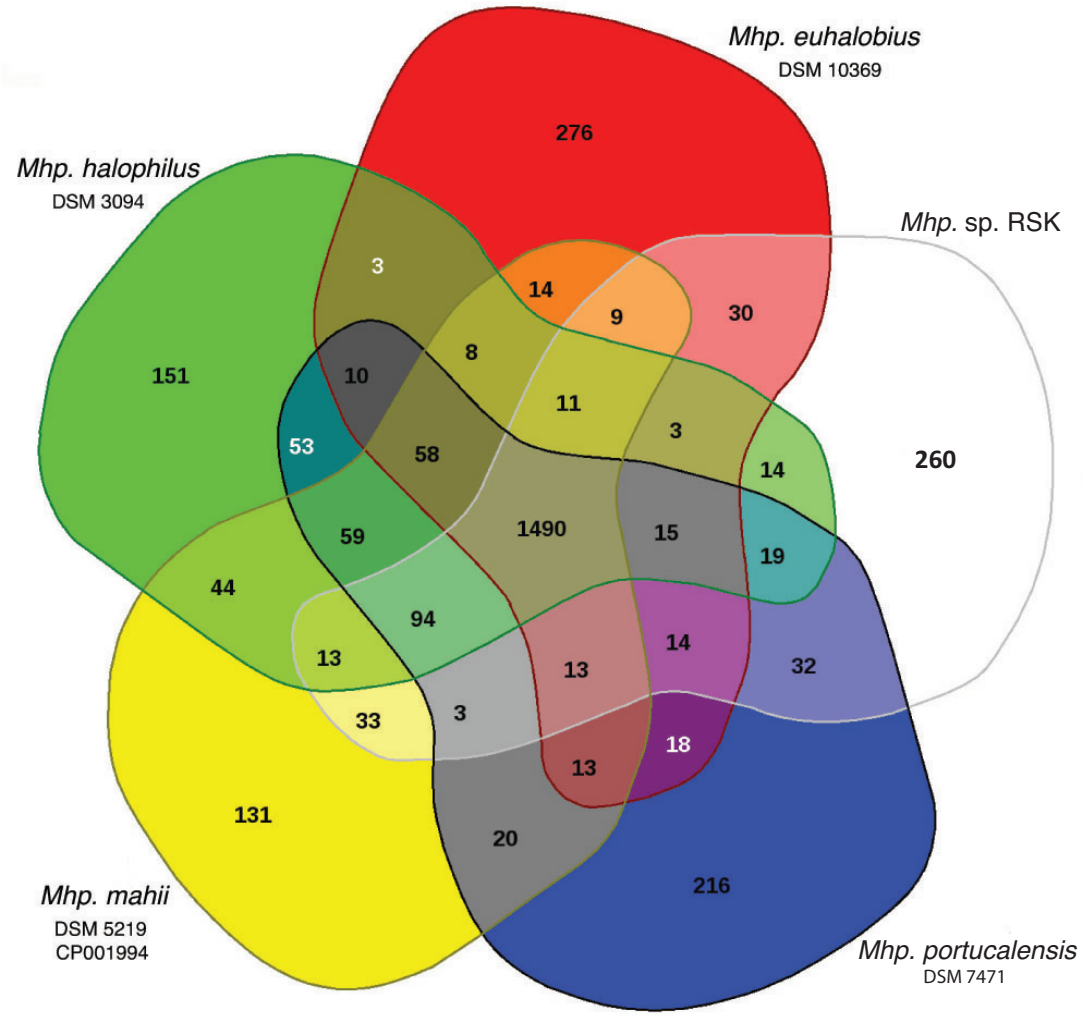


Figure 20. Comparative analysis of the predicted proteomes of *Methanohalophilus* species.

3.3.4 Core metabolic features of Methanohalophilus

In general, all newly sequenced *Methanohalophilus* genomes share very similar genomic capacities with the described complete genome of this genus — *Mehtanohalophilus mahii* (Spring *et al.*, 2010), in particular with respect to energy metabolism and to strategies for osmotic and oxidative stress response.

Methanogenesis

Methanogenesis is the major energy producing process in methanogens. Each *Methanohalophilus* genome encodes complete sets of enzymes for methanogenesis by utilizing methyl compounds. The advantages for *Methanohalophilus* spp. being methyl compound utilizing methanogens are: 1) the metabolism of methyl compounds and lack of hydrogenase prevented *Methanohalophilus* directly transfering electrons to sulfate-reducers. Previous studies revealed that when a versatile methanogen co-existed with sulfate-reducers, the methanogen benefited sulfate-reducing species by direct electron transfer (Phelps *et al.*, 1985). This process, in turn, has a negative effect on the methane yield, and, more importantly the growth (Phelps *et al.*, 1985; Allen *et al.*, 2009). Thus, methanogens which lack both hydrogenase and obligatory methylotrophy have a growth advantage in competitive environments (Allen *et al.*, 2009). 2) Not sharing common substrates such as H_2 and acetate, which methanogens generally have a lower affinity to, avoids competition with sulfate-reducers (McGenity, 2010). 3) Methyl compounds are important methane precursors and could be better substrates for methanogenesis in hypersaline conditions. In their hypersaline niche, methyl compounds are derived from salinity-induced compatible solutes and are produced by the degradation of organic compounds. They can therefore serve as direct substrate for methylotrophic methanogens (King, 2009).

Oxidative Adaptation

Methanogens inhabit a diverse array of ecosystems and a broad range of physicochemical conditions. Oxygen is one of the strongest barriers against

their distribution since they are strict anaerobes (Hoehler *et al.*, 2010). It is thus interesting that *Methanohalophilus* populations are commonly found in brine-seawater interfaces (BSIs) (Yakimov *et al.*, 2013; La Cono *et al.*, 2011; Yakimov *et al.*, 2014; Guan *et al.*, 2015) and that strain RSK was isolated from a BSI in Kebrit deep with 4.6µM oxygen (Ngugi *et al.*, 2014).

Evidently, members of *Methanohalophilus* have developed multiple mechanisms to protect them from oxygen toxicity by consuming O_2 and reactive oxygen species (ROS) on the one hand, and to even repair damaged cellular compounds and enzymes on the other. The genomes of *Methanohalophilus* encode up to 8 copies of thioredoxin and up to 6 copies of different isozymes of catalase peroxidase peroxiredoxin, alkyl-hydroperoxide reductase and alkyl-hydroperoxidase. Superoxide dismutase, catalase and superoxide reductase were also detected in each genome.

Osmotic Adaptation

Osmotic adaptation is one of the main factors that allow halophilic methanogens to thrive in hypersaline environments. All sequenced *Methanohalophilus* have the genome capacity to perform both the “salt-in” and “compatible solute” strategies (Oren, 2011) to combat osmo-stress as described for *Mhp. mahii* (Spring *et al.*, 2010). Notably, they can perform *de novo* synthesis of both betaine (which, in methanogens, have only been detected in the genus Methanohalobium, see also Chapter II) and N^{ε}-acetyl-β-lysine (which was also detected in the genome of closely related halophilic methanogens of genus *Methanococcoides* and several other methanogens.)

3.3.5 Flexible gene sets of Methanohalophilus

Sulfur Metabolism

Strain RSK is the only methanogen of genus *Methanohalophilus* to have sulfate adenylyltransferase, 3'-phosphoadenosine 5'-phosphosulfate synthase (adenylylsulfate kinase) and phosphoadenosine phosphosulfate reductase to convert sulfate first to adenylylsulfate (APS), then to 3'-phosphoadenylylsulfate (PAPS), and to sulfite for the assimilatory sulfate reduction pathway. All four other draft genomes of *Methanohalophilus* species lack 3'-phosphoadenosine 5'-phosphosulfate synthase (adenylylsulfate kinase). Sulfate adenylyltransferase was not detected in the draft genomes of *Mhp. portucalensis* or *Mhp. euhalobius*.

Homologous genes of the novel coenzyme F_{420}-dependent sulfite reductase (Fsr) (Johnson and Mukhopadhyay, 2005) have been identified in *Methanohalophilus,* species *mahii*, *halophilus*, and *portucalensis*, however not in the draft genomes of *Mhp. euhalobius* and strain RSK. The enzyme product of this gene has a dissimilatory-type siroheme sulfite reductase domain A (DsrA) and performs H_2F_{420} dependent sulfite reduction. All *Methanohalophilus* have one or two additional copies of DsrA domain containing nitrite and sulfite reductase gene (anaerobic sulfite reductase). Sulfite is toxic to methanogens in that it inactivates the essential enzyme for methanogenesis, methyl-coenzyme M reductase (Mahlert *et al.*, 2002). The sulfite reductase could thus help protect halophilic methanogens by reducing sulfite and by also contributing to the sulfite assimilation process.

ABC Transporters

Members of *Methanohalophilus* appear to have different preferences for amino acid and peptide uptake. Operons for peptide transport (*opp*) are found only in *Mhp. mahii*, *Mhp. halophilus,* and strain RSK, indicating that these members could rely on this approach to directly uptake peptide as a source of nutrient.

The livK genes involved in high-affinity branched-chain amino acid transport system are only present in the genomes of *Mhp. mahii*, *Mhp. portucalensis*, *and Mhp. halophilus* (*Mhp. euhalobius* has only a truncated version of this gene). This periplasmic Leu/Ile/Val-binding protein serves as the primary receptor for the leucine transport system in *E. coli* (Magnusson *et al.*, 2004; Anderson and Oxender, 1977). This suggests that different members of *Methanohalophilus* may have different abilities to accumulate branched-chain amino acid from the environment.

Glycogen Synthesis and Utilization

Glycogen is a common reserve polysaccharide in methanogens (Simpson and Whitman, 1993). Enzymes for glycogen metabolism (glycogen synthase, alpha-amylase, and glycogen debranching protein) are present in *Mhp. mahii* and *Mhp. halophilus,* suggesting that only these two members of *Methanohalophilus* are able to synthesize glycogen and depend on them as a carbon source during starvation.

Viral Defense

Clustered regularly interspaced short palindromic repeats (CRISPR)-Cas systems provide prokaryotes with acquired, adaptive, and heritable resistance

against foreign invading genetic elements such as viruses, phages and plasmids (Barrangou *et al.*, 2007; Horvath and Barrangou, 2010). Viral defense was not previously shown in *Methanohalophilus* because the only published *Methanohalophilus* genome (species *mahii*) has stand-alone *cas1* and *cas4* genes but no CRISPR locus. Though revealing the genomic information of the additional four genomes within this genus, we found that three out of the five *Methanohalophilus* species possess CRISPR/Cas systems as depicted in Figure 21.

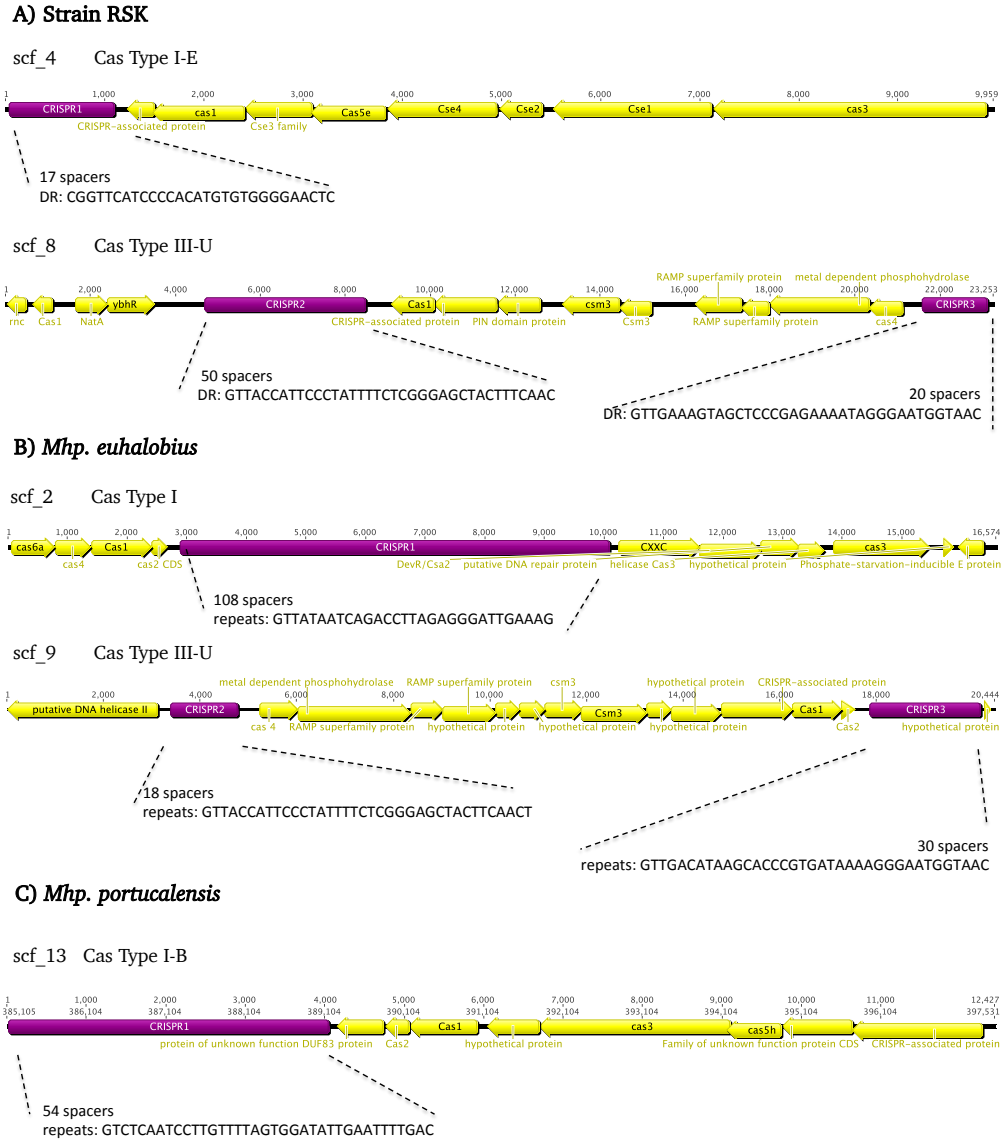


Figure 21. CRISPR/Cas systems in *Methanohalophilus*.

Both genomes of strain RSK and *Mhp. euhalobius* carry two different Cas arrays—one belonging to Type I and the other one belonging to Type III-U. Both Type III-U Cas arrays exhibit high sequence similarities and high synteny (*cas6 – cas10 – cas10/csm1 – csx10* – a gene for hypothetical protein, which might be a response regulator GlrR – hypothetical protein 2 – *csm3 – csm3 – csx1 – cas1*). Similar Cas arrays could only be found in the genomes of *Candidatus* Methanoperedens nitroreducens and *Methanospirillum hungatei* so far, suggesting a novel variant of subtype III-U. This further indicates the uniqueness of this subtype in anaerobic methane-producing and -utilizing microbes. Strain RSK's CRISPR locus 2 and *Mhp. euhalobius'* CRISPR locus 2 share the same direct repeat element (GTTACCATTCCCTATTTTCTCGGGAGCTACTTTCAAC) (Table 16). This suggests the close evolutionary relationship in viral defense systems between strain RSK and *Mhp. euhalobius*.

Table 16. CRISPR direct repeats (DRs) in *Methanohalophilus*.

Organism	CRISPR DR elements	Length (bp)	Nr. of spacers per locus	Size range of spacers (bp)	Class assignments	Structure motifs
Strain *RSK*	CGGTTCATCCCCACATGTGTGGGGAACTC*	29	28	32		motif 13
	GTTACCATTCCCTATTTTCTCGGGAGCTACTTTCAAC	37	50	34-47		
	GTTGAAAGTAGCTCCCGAGAAAATAGGGAATGGTAAC	37	20	35-43	superclass E	
Mhp. euhalobius	GTTATAATCAGACCTTAGAGGGATTGAAAG		108	33-45	superclass A	
	GTTACCATTCCCTATTTTCTCGGGAGCTACTTCAACT	37	18	35-45		
	GTTGACATAAGCACCCGTGATAAAAGGGAATGGTAAC	37	30	34-42	superclass E	

Mhp. portucalensis	GTCTCAATCCTTGTTTTAGTGGATATTGAATTTTGAC	37	54	32-47

other questionable CRISPR DR elements

Mhp. RSK	ATCTTCTCCTGTCATTGTGGTTGGCAGTTCTGTCCAT	37	1	35
	AATACAGGTACTATAAATCAAAGTTATTCGAC	32	1	49
Mhp. portucalensis	GGATAGCACTTTGACTGATAATTC	24	1	42
Mhp. mahii	GGAGTCTATCTCTATAGTTCAAGT		1	45
Mhp. halophilus	TCATCTGCATATCCATCTCTATC		1	43

*With the exception of the first DR in the locus as TGTTTCATCCCCACATGTGTGGGGAACTC.

3.3.6 Unique genes in strain RSK

The 260 predicted species-specific genes in the genome of RSK span 15 COG categories and are overrepresented in replication, recombination and repair (22 genes), cell wall/membrane/envelope biogenesis (17 genes), inorganic ion transport and metabolism (14 genes), general function prediction only (14 genes), and coenzyme transport and metabolism (12 genes) (also see Table S1). The unique gene composition of this strain indicates that this brine *Mehtanohalophilus* could be capable of using a range of cellular functions to cope with the deep-sea hypersaline habitat. For example, unique membrane characteristics (reflected by abundant unique glycosyltransferases and surface proteins that are associated with cell surface compositions, multiple species-specific transporters that are responsible for taking in useful substances and removing toxic compounds), a putative ability to catalyze ethylbenzene degradation (reflected by its predicted ethylbenzene dehydrogenase), and a reinforced transcription regulation apparatus under extreme conditions (reflected by multiple unique copies of HxlR and ArsR family transcriptional regulators).

3.4 Conclusions

In this chapter, we described the isolation and characterization of a novel species of genus *Methanohalophilus* obtained from Kebrit Deep in the Red Sea. We generated genome sequences of multiple, closely related *Methanohalophilus* species and we conducted a comparative genomics analysis on this genus. We have elaborated on CRISPR/Cas system, transporters, and carbon hydrate metabolism, as some of the most important factors that help different *Methanohalophilus* populations to adapt and colonize various hypersaline habitats. This chapter has laid a solid foundation for the following in-depth and comprehensive genomic analysis of this genus.

3.5 Supplementary materials

Table S1. Unique genes in strain RSK.

Gene ID	INDIGO annotation	COG classification
RSK_00039	Low-affinity putrescine importer PlaP	Amino acid transport and metabolism
RSK_00417	Glycine betaine-L-proline transport system permease protein ProW	Amino acid transport and metabolism
RSK_00491	putative cysteine desulfurase	Amino acid transport and metabolism
RSK_00642	Arginine-ornithine antiporter	Amino acid transport and metabolism
RSK_00834	Sodium-proline symporter	Amino acid transport and metabolism
RSK_01055	Osmoregulated proline transporter	Amino acid transport and metabolism
RSK_01141	putative acetolactate synthase large subunit	Amino acid transport and metabolism
RSK_01194	shikimate kinase	Amino acid transport and metabolism
RSK_02024	Oligopeptide transport ATP-binding protein OppF	Amino acid transport and metabolism
RSK_02028	Glutathione-binding protein GsiB	Amino acid transport and metabolism
RSK_01214	Phosphoenolpyruvate synthase	Carbohydrate transport and metabolism
RSK_01277	Phosphoenolpyruvate synthase	Carbohydrate transport and metabolism
RSK_01503	Aquaporin AqpM	Carbohydrate transport and metabolism
RSK_02209	Putative membrane-bound acyltransferase YfiQ	Carbohydrate transport and metabolism
RSK_00111	UDP-glucose-GDP-mannose dehydrogenase	Cell wall/membrane/envelope biogenesis
RSK_00124	D-inositol 3-phosphate glycosyltransferase	Cell wall/membrane/envelope

		biogenesis
RSK_00125	GDP-mannose-dependent alpha--phosphatidylinositol mannosyltransferase	Cell wall/membrane/envelope biogenesis
RSK_00127	Putative UDP-glucose 4-epimerase	Cell wall/membrane/envelope biogenesis
RSK_00128	Capsular polysaccharide biosynthesis glycosyltransferase CapM	Cell wall/membrane/envelope biogenesis
RSK_00129	hexosyltransferase	Cell wall/membrane/envelope biogenesis
RSK_00130	PGA biosynthesis protein CapA	Cell wall/membrane/envelope biogenesis
RSK_00132	UDP-N-acetylglucosamine 2-epimerase	Cell wall/membrane/envelope biogenesis
RSK_00135	Undecaprenyl-phosphate 4-deoxy-4-formamido-L-arabinose transferase	Cell wall/membrane/envelope biogenesis
RSK_00166	glycosyl transferase group 1 family protein	Cell wall/membrane/envelope biogenesis
RSK_00168	Protein RfbU	Cell wall/membrane/envelope biogenesis
RSK_00169	GDP-mannose 46-dehydratase	Cell wall/membrane/envelope biogenesis
RSK_00171	Glycosyl transferase group 1	Cell wall/membrane/envelope biogenesis
RSK_00174	D-inositol 3-phosphate glycosyltransferase	Cell wall/membrane/envelope biogenesis
RSK_00176	1-deoxy-D-xylulose-5-phosphate synthase	Cell wall/membrane/envelope biogenesis
RSK_00473	MscS Mechanosensitive ion channel	Cell wall/membrane/envelope biogenesis
RSK_01014	S-layer domain protein	Cell wall/membrane/envelope biogenesis
RSK_00048	Ribose 15-bisphosphate isomerase	Coenzyme transport and metabolism
RSK_00131	Indoleacetate--lysine synthetase	Coenzyme transport and metabolism
RSK_00445	phenylacetate--CoA ligase	Coenzyme transport and metabolism
RSK_00446	Cyclic pyranopterin monophosphate synthase	Coenzyme transport and metabolism
RSK_00448	Methylated-thiol--coenzyme M methyltransferase	Coenzyme transport and metabolism
RSK_00451	Biotin synthase	Coenzyme transport and metabolism
RSK_01139	Malonyl-acyl-carrier protein O-methyltransferase	Coenzyme transport and metabolism
RSK_01314	Uroporphyrinogen-III decarboxylase Coenzyme metabolism	Coenzyme transport and metabolism
RSK_01386	5-amino-6-uracil reductase	Coenzyme transport and metabolism
RSK_01632	von Willebrand factor type A	Coenzyme transport and metabolism
RSK_01678	uroporphyrinogen decarboxylase	Coenzyme transport and metabolism
RSK_02029	Demethylmenaquinone methyltransferase	Coenzyme transport and metabolism
RSK_00550	CRISPR-associated Cse3 family protein	CONSERVED DOMAIN-CD
RSK_00551	CRISPR-associated Cas5e family protein	CONSERVED DOMAIN-CD
RSK_00167	type 11 methyltransferase	CONSERVED DOMAIN-PFAM
RSK_00552	CRISPR-associated Cse4 family protein	CONSERVED DOMAIN-PFAM
RSK_00613	ArsR family transcriptional regulator	CONSERVED DOMAIN-PFAM
RSK_00914	Zinc metalloprotease ZmpB	CONSERVED DOMAIN-PFAM

RSK_02347	transposase IS605 OrfB family protein	CONSERVED DOMAIN-PFAM
RSK_02201	cobalt transport protein CbiN	CONSERVED DOMAIN-PRK
RSK_00463	4Fe-4S ferredoxin iron-sulfur binding domain protein	CONSERVED DOMAIN-TIGR
RSK_00915	PKD domain containing protein	CONSERVED DOMAIN-TIGR
RSK_01229	peptidase S8	CONSERVED DOMAIN-TIGR
RSK_01836	PKD domain containing protein	CONSERVED DOMAIN-TIGR
RSK_01934	Dimethylamine methyltransferase MtbB2	CONSERVED DOMAIN-TIGR
RSK_00201	Nod factor export ATP-binding protein I	Defense mechanisms
RSK_01015	Macrolide export ATP-binding-permease protein MacB	Defense mechanisms
RSK_01593	Type-1 restriction enzyme StySJI specificity protein	Defense mechanisms
RSK_01789	Inner membrane transport permease YbhR	Defense mechanisms
RSK_00060	formylmethanofuran dehydrogenase subunit E	Energy production and conversion
RSK_00061	formylmethanofuran dehydrogenase subunit E	Energy production and conversion
RSK_00197	CoB--CoM heterodisulfide reductase 2 iron-sulfur subunit D	Energy production and conversion
RSK_00204	phosphoadenosine phosphosulfate reductase	Energy production and conversion
RSK_00455	NADH-dependent flavin oxidoreductase	Energy production and conversion
RSK_00457	putative aldehyde dehydrogenase AldX	Energy production and conversion
RSK_01788	ATP-binding transport protein NatA	Energy production and conversion
RSK_00601	Bifunctional enzyme fae-hp	Function unknown
RSK_01313	Inner membrane protein YaaH	Function unknown
RSK_00062	putative permease General function prediction only	General function prediction only
RSK_00091	ATPase	General function prediction only
RSK_00158	PKD domain containing protein	General function prediction only
RSK_00163	Putative O-antigen transporter	General function prediction only
RSK_00172	sugar transferase	General function prediction only
RSK_00459	Serine-threonine-protein kinase pkn1	General function prediction only
RSK_00464	Dimethylamine corrinoid protein 2	General function prediction only
RSK_00555	Putative CRISPR-associated nuclease-helicase Cas3	General function prediction only
RSK_00638	Dimethylamine corrinoid protein 1	General function prediction only
RSK_00829	metallo-beta-lactamase superfamily hydrolase	General function prediction only
RSK_01315	Dimethylamine corrinoid protein 2	General function prediction only
RSK_01344	methyltransferase	General function prediction only
RSK_01675	47 kDa protein	General function prediction only
RSK_01837	Immunoglobulin A1 protease	General function prediction only
RSK_00036	cell surface protein	Inorganic ion transport and metabolism
RSK_00038	H+-transporting ATPase	Inorganic ion transport and metabolism
RSK_00393	Poly C5 epimerase	Inorganic ion transport and metabolism
RSK_00452	Putative binding protein YtlA	Inorganic ion transport and metabolism
RSK_00453	Binding-protein-dependent transport system inner membrane component protein	Inorganic ion transport and metabolism
RSK_00454	Aliphatic sulfonates import ATP-binding protein SsuB	Inorganic ion transport and metabolism
RSK_00576	Ferrous iron transport protein B	Inorganic ion transport and metabolism
RSK_01211	putative cation-transporting ATPase F	Inorganic ion transport and metabolism
RSK_01501	Arsenite resistance protein ArsB	Inorganic ion transport and metabolism

RSK_01882	Poly C5 epimerase	Inorganic ion transport and metabolism
RSK_02199	ABC transporter protein	Inorganic ion transport and metabolism
RSK_02200	Energy-coupling factor transporter transmembrane protein EcfT	Inorganic ion transport and metabolism
RSK_02202	Cobalt uptake substrate-specific transmembrane region protein	Inorganic ion transport and metabolism
RSK_02300	Catalase-peroxidase	Inorganic ion transport and metabolism
RSK_00458	3-oxoacyl-acyl-carrier-protein reductase FabG	Lipid transport and metabolism
RSK_01152	O-linked N-acetylglucosamine transferase	Posttranslational modification, protein turnover, chaperones
RSK_01781	Lon protease	Posttranslational modification, protein turnover, chaperones
RSK_02057	peptidase U32	Posttranslational modification, protein turnover, chaperones
RSK_00077	tRNA -dimethyltransferase	Replication, recombination and repair
RSK_00121	DNA replication	Replication, recombination and repair
RSK_00237	Transposase	Replication, recombination and repair
RSK_00238	Putative transposase in snaA-snaB intergenic region	Replication, recombination and repair
RSK_00262	Modification methylase SmaI	Replication, recombination and repair
RSK_00392	transposase	Replication, recombination and repair
RSK_00523	Putative transposase in snaA-snaB intergenic region	Replication, recombination and repair
RSK_00549	CRISPR-associated endonuclease Cas1	Replication, recombination and repair
RSK_00620	Putative transposase in snaA-snaB intergenic region	Replication, recombination and repair
RSK_00670	transposase	Replication, recombination and repair
RSK_00852	nicotinate-nucleotide--dimethylbenzimidazole phosphoribosyltransferase	Replication, recombination and repair
RSK_01257	DNA double-strand break repair Rad50 ATPase	Replication, recombination and repair
RSK_01259	ATP-dependent RNA helicase RhlB//DNA repair	Replication, recombination and repair
RSK_01260	Modification methylase Cfr9I	Replication, recombination and repair
RSK_01289	PHP domain protein	Replication, recombination and repair
RSK_01316	transposase	Replication, recombination and repair
RSK_01385	UvrABC system protein A	Replication, recombination and repair
RSK_01428	Putative transposase in snaA-snaB intergenic region	Replication, recombination and repair
RSK_01774	hypothetical protein	Replication, recombination and repair
RSK_01778	putative modification methylase HindII	Replication, recombination and repair
RSK_01981	Transposase	Replication, recombination and repair
RSK_02348	Putative transposase in snaA-snaB intergenic region	Replication, recombination and repair
RSK_00084	Signal transduction histidine-protein kinase AtoS	Signal transduction mechanisms
RSK_00461	Sensor histidine kinase ResE	Signal transduction mechanisms
RSK_00643	Universal stress protein	Signal transduction mechanisms
RSK_01091	Sensor histidine kinase ResE	Signal transduction mechanisms
RSK_01109	Polar-differentiation response regulator DivK	Signal transduction mechanisms
RSK_01110	multi-sensor signal transduction histidine kinase	Signal transduction mechanisms
RSK_01212	signal transduction histidine kinase	Signal transduction mechanisms
RSK_01296	PAS-PAC sensor protein	Signal transduction mechanisms
RSK_01623	multi-sensor signal transduction histidine kinase	Signal transduction mechanisms

RSK_00199	HxlR family transcriptional regulator	Transcription
RSK_00205	ArsR family transcriptional regulator	Transcription
RSK_00460	ArsR family transcriptional regulator	Transcription
RSK_00600	HxlR family transcriptional regulator	Transcription
RSK_00993	DNA-directed RNA polymerase subunit P	Transcription
RSK_01318	ArsR family transcriptional regulator	Transcription
RSK_01387	MarR family transcriptional regulator	Transcription
RSK_01785	Ribonuclease 3	Transcription
RSK_00104	Putative toxin RelE4	Translation, ribosomal structure and biogenesis
RSK_00120	Lysine--tRNA ligase	Translation, ribosomal structure and biogenesis
RSK_00001	hypothetical protein	
RSK_00037	cell surface protein	
RSK_00096	Archaeal ATPase protein	
RSK_00098	hypothetical protein	
RSK_00109	hypothetical protein	
RSK_00112	hypothetical protein	
RSK_00126	hypothetical protein	
RSK_00134	hypothetical protein	
RSK_00138	hypothetical protein	
RSK_00147	hypothetical protein	
RSK_00160	hypothetical protein	
RSK_00196	hypothetical protein	
RSK_00200	hypothetical protein	
RSK_00206	Peptidase C39 like protein	
RSK_00222	SagB-type dehydrogenase domain protein	
RSK_00253	hypothetical protein	
RSK_00254	hypothetical protein	
RSK_00255	hypothetical protein	
RSK_00256	hypothetical protein	
RSK_00257	hypothetical protein	
RSK_00258	Carboxypeptidase regulatory-like domain protein	
RSK_00259	hypothetical protein	
RSK_00260	hypothetical protein	
RSK_00261	hypothetical protein	
RSK_00263	hypothetical protein	
RSK_00264	hypothetical protein	
RSK_00265	hypothetical protein	
RSK_00266	hypothetical protein	
RSK_00267	HTH domain protein	
RSK_00268	hypothetical protein	
RSK_00269	hypothetical protein	
RSK_00286	Epoxyqueuosine reductase	
RSK_00358	NYN domain protein	

RSK_00381	hypothetical protein
RSK_00391	transposase
RSK_00394	hypothetical protein
RSK_00415	hypothetical protein
RSK_00447	hypothetical protein
RSK_00449	hypothetical protein
RSK_00456	hypothetical protein
RSK_00462	Blue-light-activated histidine kinase
RSK_00465	radical SAM domain-containing protein
RSK_00472	MscS Mechanosensitive ion channel
RSK_00547	hypothetical protein
RSK_00548	CRISPR-associated protein
RSK_00553	CRISPR-associated Cse2 protein
RSK_00554	CRISPR-associated Cse1 protein
RSK_00593	hypothetical protein
RSK_00602	hypothetical protein
RSK_00608	Nuclease-related domain protein
RSK_00671	hypothetical protein
RSK_00672	Ethylbenzene dehydrogenase protein
RSK_00673	hypothetical protein
RSK_00726	hypothetical protein
RSK_00744	hypothetical protein
RSK_00826	hypothetical protein
RSK_00828	hypothetical protein
RSK_00830	hypothetical protein
RSK_00835	hypothetical protein
RSK_00870	hypothetical protein
RSK_00876	Winged helix-turn-helix DNA-binding protein
RSK_00880	Thioredoxin domain protein
RSK_00891	hypothetical protein
RSK_00952	phosphate uptake regulator PhoU
RSK_01016	MacB-like periplasmic core domain protein
RSK_01079	hypothetical protein
RSK_01090	diguanylate cyclase
RSK_01112	hypothetical protein
RSK_01131	hypothetical protein
RSK_01151	Restriction endonuclease protein
RSK_01153	hypothetical protein
RSK_01171	hypothetical protein
RSK_01187	hypothetical protein
RSK_01256	hypothetical protein
RSK_01258	hypothetical protein
RSK_01266	hypothetical protein
RSK_01275	ATPase AAA

RSK_01279	pyridoxamine 5'-phosphate oxidase-related FMN-binding protein
RSK_01288	hypothetical protein
RSK_01300	hypothetical protein
RSK_01305	hypothetical protein
RSK_01310	hypothetical protein
RSK_01311	hypothetical protein
RSK_01312	GCN5-related N-acetyltransferase
RSK_01317	transposase
RSK_01322	Reprolysin domain containing protein
RSK_01393	hypothetical protein
RSK_01480	hypothetical protein
RSK_01502	hypothetical protein
RSK_01588	hypothetical protein
RSK_01595	hypothetical protein
RSK_01624	hypothetical protein
RSK_01633	von Willebrand factor type A domain protein
RSK_01681	PAS-PAC sensor signal transduction histidine kinase
RSK_01728	conserved repeat domain protein
RSK_01773	hypothetical protein
RSK_01775	hypothetical protein
RSK_01776	hypothetical protein
RSK_01777	ATP-dependent DNA helicase RecG
RSK_01779	hypothetical protein
RSK_01780	DNA repair protein
RSK_01782	hypothetical protein
RSK_01783	hypothetical protein
RSK_01787	hypothetical protein
RSK_01790	hypothetical protein
RSK_01791	hypothetical protein
RSK_01792	hypothetical protein
RSK_01793	hypothetical protein
RSK_01807	hypothetical protein
RSK_01973	hypothetical protein
RSK_01974	hypothetical protein
RSK_02036	sulfopyruvate decarboxylase subunit beta
RSK_02067	hypothetical protein
RSK_02132	hypothetical protein
RSK_02191	Metallo-peptidase domain containing protein
RSK_02195	Metallo-peptidase domain containing protein
RSK_02210	Acyltransferase domain containing protein
RSK_02221	Winged helix-turn-helix DNA-binding protein

3.6 References

Alam I, Antunes A, Kamau AA, alawi WB, Kalkatawi M, Stingl U, *et al.* (2013). INDIGO – INtegrated Data Warehouse of MIcrobial GenOmes with examples from the Red Sea extremophiles Hernandez-Lemus, E (ed). *PLoS ONE* **8**: e82210.

Allen MA, Lauro FM, Williams TJ, Burg D, Siddiqui KS, De Francisci D, *et al.* (2009). The genome sequence of the psychrophilic archaeon, *Methanococcoides burtonii*: the role of genome evolution in cold adaptation. *ISME J* **3**: 1012–1035.

Anderson JJ, Oxender DL. (1977). *Escherichia coli* transport mutants lacking binding protein and other components of the branched-chain amino acid transport systems. **130**: 384–392.

Antunes A, Ngugi DK, Stingl U. (2011). Microbiology of the Red Sea (and other) deep-sea anoxic brine lakes. *Environmental Microbiology Reports* **3**: 416–433.

Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, *et al.* (2012). SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol* **19**: 455–477.

Barrangou R, Fremaux C, Deveau H, Richards M, Boyaval P, Moineau S, *et al.* (2007). CRISPR provides acquired resistance against viruses in prokaryotes. *Science* **315**: 1712.

Bezrukova LV, Belyaev SS. (1987). Properties of the coccoid methylotrophic methanogen, *Methanococcoides euhalobius* sp. *Mikrobiologiia* **56**: 523–527.

Blom J, Albaum SP, Doppmeier D, Pühler A, Vorhölter F-J, Zakrzewski M, *et al.* (2009). EDGAR: a software framework for the comparative analysis of prokaryotic genomes. *BMC Bioinf* **10**: 154.

Boone DR, Mathrani IM, Liu Y, Menaia JAGF, Mah RA, Boone JE. (1993). Isolation and characterization of *Methanohalophilus portucalensis* sp. nov. and DNA reassociation study of the genus *Methanohalophilus*. *Int J Syst Bacteriol* **43**: 430–437.

Cono V, Arcadi E, Spada G, Barreca D, Laganà G, Bellocco E, *et al.* (2015). A three-component microbial consortium from deep-sea salt-saturated anoxic lake Thetis links anaerobic glycine betaine degradation with methanogenesis. *Microorganisms 2015, Vol 3, Pages 500-517* **3**: 500–517.

Eder W, Jahnke LL, Schmidt M, Huber R. (2001). Microbial diversity of the brine-seawater interface of the Kebrit Deep, Red Sea, studied via 16S rRNA gene sequences and cultivation methods. *Appl Environ Microbiol* **67**: 3077–3085.

Goris J, Konstantinidis KT, Klappenbach JA, Coenye T, Vandamme P, Tiedje JM. (2007). DNA-DNA hybridization values and their relationship to whole-genome sequence similarities. *Int J Syst Evol Microbiol* **57**: 81–91.

Guan Y, Hikmawan T, Antunes A, Ngugi D, Stingl U. (2015). Diversity of methanogens and sulfate-reducing bacteria in the interfaces of five deep-sea anoxic brines of the Red Sea. *Research in Microbiology*. doi:10.1016/j.resmic.2015.07.002.

Hoehler T, Gunsalus RP, McInerney MJ. (2010). Environmental constraints that limit methanogenesis. In: *Handbook of Hydrocarbon and Lipid Microbiology*, Springer Berlin Heidelberg: Berlin, Heidelberg, pp 635–654.

Horvath P, Barrangou R. (2010). CRISPR/Cas, the immune system of bacteria and archaea. *Science* **327**: 167–170.

Johnson EF, Mukhopadhyay B. (2005). A new type of sulfite reductase, a novel coenzyme F420-dependent enzyme, from the methanarchaeon *Methanocaldococcus jannaschii*. *J Biol*

Chem **280**: 38776–38786.

Katayama T, Yoshioka H, Mochimaru H, Meng X-Y, Muramoto Y, Usami J, *et al.* (2014). *Methanohalophilus levihalophilus* sp. nov., a slightly halophilic, methylotrophic methanogen isolated from natural gas-bearing deep aquifers, and an emended description of the genus *Methanohalophilus*. *Int J Syst Evol Microbiol* **64**: ijs.0.063677–0–2093.

King GM. (2009). Utilization of hydrogen, acetate, and 'noncompetitive'; substrates by methanogenic bacteria in marine sediments. *Geomicrobiology Journal* **3**: 275–306.

La Cono V, Smedile F, Bortoluzzi G, Arcadi E, Maimone G, Messina E, *et al.* (2011). Unveiling microbial life in new deep-sea hypersaline Lake Thetis. Part I: Prokaryotes and environmental settings. *Environ Microbiol* **13**: 2250–2268.

Magnusson U, Salopek-Sondi B, Luck LA, Mowbray SL. (2004). X-ray structures of the leucine-binding protein illustrate conformational changes and the basis of ligand specificity. *J Biol Chem* **279**: 8747–8752.

Mahlert F, Bauer C, Jaun B, Thauer RK, Duin EC. (2002). The nickel enzyme methyl-coenzyme M reductase from methanogenic archaea: In vitro induction of the nickel-based MCR-ox EPR signals from MCR-red2. *J Biol Inorg Chem* **7**: 500–513.

McGenity TJ. (2010). Methanogens and methanogenesis in hypersaline environments. In: *Handbook of Hydrocarbon and Lipid Microbiology,* Springer Berlin Heidelberg: Berlin, Heidelberg, pp 665–680.

Ngugi DK, Blom J, Alam I, Rashid M, Ba-Alawi W, Zhang G, *et al.* (2014). Comparative genomics reveals adaptations of a halotolerant thaumarchaeon in the interfaces of brine pools in the Red Sea. *ISME J*. doi:10.1038/ismej.2014.137.

Oren A. (2011). Thermodynamic limits to microbial life at high salt concentrations. *Environ Microbiol* **13**: 1908–1923.

Paterek JR, Smith PH. (1988). *Methanohalophilus mahii* gen nov., sp. nov., a methylotrophic halophilic methanogen. *Int J Syst Bacteriol* **38**: 122–123.

Phelps TJ, Conrad R, Zeikus JG. (1985). Sulfate-dependent interspecies H(2) Transfer between *Methanosarcina barkeri* and *Desulfovibrio vulgaris* during coculture metabolism of acetate or methanol. *Appl Environ Microbiol* **50**: 589–594.

Richter M, Rosselló-Mora R. (2009). Shifting the genomic gold standard for the prokaryotic species definition. *Proc Natl Acad Sci USA* **106**: 19126–19131.

Simpson PG, Whitman WB. (1993). Anabolic pathways in methanogens. In: *Methanogenesis*, Ferry, JG (ed), Springer US: Boston, MA, pp 445–472.

Sowers KR, Baron SF, Ferry JG. (1984). *Methanosarcina acetivorans* sp. nov., an acetotrophic methane-producing bacterium isolated from marine sediments. *Appl Environ Microbiol* **47**: 971–978.

Sowers KR, Ferry JG. (1983). Isolation and characterization of a methylotrophic marine methanogen, *Methanococcoides methylutens* gen. nov., sp. nov. *Appl Environ Microbiol* **45**: 684–690.

Spring S, Scheuner C, Lapidus A, Lucas S, Glavina Del Rio T, Tice H, *et al.* (2010). The genome sequence of *Methanohalophilus mahii* SLP(T) reveals differences in the energy metabolism among members of the *Methanosarcinaceae* inhabiting freshwater and saline environments. *Archaea* **2010**: 690737.

Thauer RK, Kaster A-K, Seedorf H, Buckel W, Hedderich R. (2008). Methanogenic archaea: ecologically relevant differences in energy conservation. *Nature Reviews Microbiology* **6**: 579–591.

van der Wielen PWJJ. (2005). The enigma of prokaryotic life in deep hypersaline anoxic basins. *Science* **307**: 121–123.

Yakimov MM, La Cono V, La Spada G, Bortoluzzi G, Messina E, Smedile F, *et al.* (2014). Microbial community of the deep-sea brine Lake Kryos seawater-brine interface is active below the chaotropicity limit of life as revealed by recovery of mRNA. *Environ Microbiol* **17**: n/a–n/a.

Yakimov MM, La Cono V, Slepak VZ, La Spada G, Arcadi E, Messina E, *et al.* (2013). Microbial life in the Lake Medee, the largest deep-sea salt-saturated formation. **3**. doi:10.1038/srep03554.

Yelton AP, Thomas BC, Simmons SL, Wilmes P, Zemla A, Thelen MP, *et al.* (2011). A semi-quantitative, synteny-based method to improve functional predictions for hypothetical and poorly annotated bacterial and archaeal genes. Eisen, JA (ed). *PLoS Comp Biol* **7**: e1002230.

Zhilina TN. (1983). A new obligate halophilic methane-producing bacterium. *Mikrobiologiia* **52**: 375–383.

4 CHAPTER IV SINGLE-CELL GENOMICS REVEALS POTENTIAL FOR PYRROLYSINE SYNTHESIS AND DECODING IN UNCULTIVATED CANDIDATE DIVISION MSBL1

Yue Guan[a], Mohamed F. Haroon[a], Intikhab Alam[b], James G. Ferry[c], Ulrich Stingl[a]

[a]Red Sea Research Center, King Abdullah University of Science and Technology, Thuwal, Saudi Arabia

[b]Computational Bioscience Research Center, King Abdullah University of Science and Technology, Thuwal, Saudi Arabia

[c]Department of Biochemistry and Molecular Biology, Pennsylvania State University, University Park, Pennsylvania, USA

Author contributions:

Yue Guan analyzed the data and wrote this chapter. Intikhab Alam maintained the bioinformatics platform for SAGs, Ulrich Stingl, James G. Ferry, and Mohamed F. Haroon provided critical comments to this chapter.

Abstract

Pyrrolysine (Pyl), the 22nd canonical amino acid, is decoded and synthesized by a very limited number of organisms in the domains of Archaea and Bacteria. It is encoded by the amber codon UAG, which is usually a stop codon. It is crucial for the functionality of methylamine methyltransferases that initiates the corrinoid-dependent methyl transfer reactions. So far all pyrrolysine-decoding organisms reported in the archaeal domain are methylotrophic methanogens. In this chapter, we report the discovery of a putative *pyl* gene cluster from single amplified genomes belonging to the archaeal Mediterranean Sea Brine Lakes group 1(MSBL1) that was recently recovered from the deep-sea brine pools of the Red Sea. Together with a complete corrinoid-dependent methyl-transfer pathway, our finding indicates that members of MSBL1 may be capable of synthesizing pyrrolysine and of metabolizing methylated compounds. This chapter provides hypothesis and insights into some unusual cases of pyrrolysine decoding and the phylogeny in the domain Archaea.

4.1 Introduction

Pyrrolysine (Pyl, O), the 22nd proteogenic amino acid, is restricted to only a few organisms and a very small number of proteins (Prat *et al.*, 2012). Pyrrolysine is crucial for methylamine metabolism (Srinivasan *et al.*, 2002; Gaston *et al.*, 2011). A common feature for substrate specific methylamine methyltransferases (MtmB, MtbB and MttB) is the incorporation of Pyl, with the in-frame and read-through amber codon (UAG) translated as Pyl. In addition, Pyl can also be found in certain proteins that are not related to methylamine metabolism as a result of neutral evolution which confers no selective advantage (Heinemann *et al.*, 2009; Prat *et al.*, 2012). Prior to this study, all known Pyl decoding archaea were phylogenetically affiliated to either the family of *Methanosarcinaceae* (Prat *et al.*, 2012), genus *Methanomassiliicoccus* (Borrel *et al.*, 2014) or *Candidatus* methanogens of the order *Methanomassiliicoccales*: *Methanomethylophilus*, and *Methanomassiliicoccus* (Borrel *et al.*, 2014), *Methanoplasma* (Lang *et al.*, 2015), *Thermoplasmatales* archaeon BRNA1. General shared features of all these members of pyrrolysine decoding archaea include minimizing their TAG codon frequencies in their genomes and possessing the gene cassette *pylTSBCD* required for biosynthesis and genetically encoding UAG codons as pyrrolysine, and constitutively encoding Pyl (Prat *et al.*, 2012; Gaston *et al.*, 2011).

In this work, we show that members of uncultivated Candidate division Mediterranean Sea Brine Lakes group 1(MSBL1) from the deep-sea brines and *Methermicoccus shengliensis*, a thermophilic and methylotrophic methanogenic archaeon that has previously been isolated from oil-production water, are

putatively Pyl decoding organisms. We also propose that members of MSBL1 could utilize and oxidize the methyl group from various methyl donors to CO_2 in the reversed CO_2 reductive pathway to produce reducing equivalents.

4.2 Methods

Genome sequences were obtained from Genbank. The secondary structure of MSBL1 $tRNA^{pyl}$ was predicted using an Aragorn tRNA scan (Laslett and Canback, 2004). Operon prediction was performed via the FGENESB online server (http://www.softberry.com/berry.phtml?topic=fgenesb&group=programs&subgroup=gfindb). Each set of protein sequences (PylB/C/D/S) were aligned by MergeAlign (Collingridge and Kelly, 2012) using MAFFT (Katoh *et al.*, 2005) and 91 other different amino acid substitution matrices. Only columns with a score ≥90% from the constituent alignments were kept as a consensus alignment. Finally, maximum-likelihood phylogenetic inference was performed based on individual protein sequences reconstructed by PHYML (Guindon *et al.*, 2010) using Le and Gascuel's (LG) (Le and Gascuel, 2008) amino acid substitution matrix. Branch support was calculated using 500 bootstraps.

4.3 Results and discussion

MSBL1 is one of the most prevalent and ubiquitous phylotypes in the deep-sea anoxic brines in both the Mediterranean Sea and the Red Sea (van der Wielen, 2005; Guan *et al.*, 2015). However, there is no cultivated MSBL1 representative to date. Recently, Mwirichia and colleagues (Mwirichia *et al.*, accepted) have reported 32 MSBL1 single-amplified genomes (SAGs) with less

than 60% estimated completeness from four deep-sea brines of the Red Sea (Genbank accession numbers LHXJ00000000-LHYO00000000). Importantly, these SAGs have been decontaminated from foreign sequences based on GC content, size, phylogenetic affiliation, and tetranucleotide frequency.

In addition to the metabolic analyses that were reported by Mwirichia *et al.,* we discovered the presence of genes responsible for catalyzing the methyltransfer reaction of methyl-compound to coenzyme M, suggesting that MSBL1 is capable of metabolizing a variety of methylamines (Table S1). Notably, we identified five Pyl-containing methylamine methyltransferases in three MSBL1 SAGs: 382A03, 382A13, and 382A20 (Table 17). Most of them have a higher similarity to methyltransferases in methanogenic species in the domain Archaea than to Bacteria by comparison against the NCBI-non-redundant protein database (Table S2).

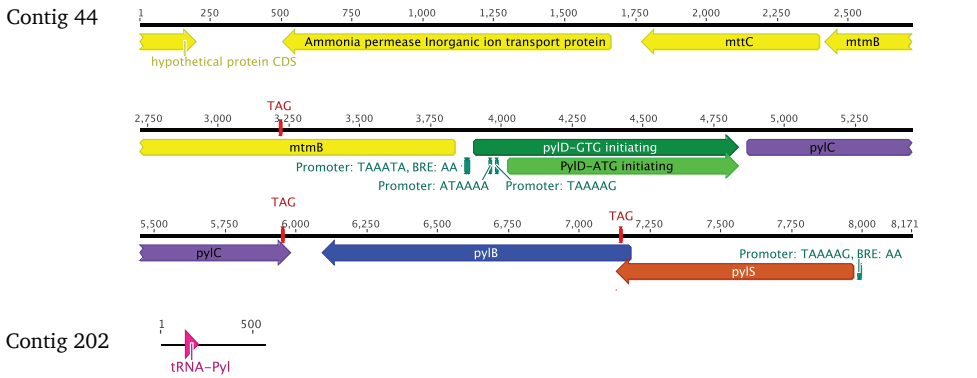
Table 17. Pyl-containing methylamine methyltransferases in the MSBL1 SAGs.

SAG ID	Contig	Coding region	Functional annotation	Protein IDs by INDIGO annotation
382A03	6	8573 -> 7197	Dimethylamine methyltransferase	AAA382A03_00124, AAA382A03_00123 and their intergenic region
	1	5402 -> 6868	Trimethylamine methyltransferase	AAA382A03_00049, AAA382A03_00050 and their intergenic region
	3	5148 -> 6587	Trimethylamine methyltransferase	AAA382A03_00087, AAA382A03_00088 and their intergenic region
382A13	31	2105 -> 678	Trimethylamine methyltransferase	AAA382A13_00401, AAA382A13_00402 and their intergenic region
382A20	44	3839 -> 2418	Monomethylamine methyltransferase	AAA382A20_00854, AAA382A20_00853 and their intergenic region

The predicted Pyl-incorporation proteins listed in Table 17 initiated the search of pyrrolysine *de novo* synthesis genes in the SAGs of MBSL1. A crucial enzyme, pyrrolysyl-tRNA synthetase (PylS), required for UAG codon suppression, is encoded in both MSBL1 SAGs 382A20 and 382A03 and validated by BLAST (Altschul *et al.*, 1990). We also identified all the genes for

the complete machinery for Pyl biosynthesis incorporation in SAG382A20 including PylBCD and a predicted $tRNA^{Pyl}$ (Figure S1) and the canonical $tRNA^{Pyl}$ synthetase (PylS). However, in SAG 382A03, the *pyl* gene cassette has been interrupted by a transposase, causing only pyrrolysine biosynthesis radical SAM protein (pylB), pyrrolysyl-tRNA synthetase and a Pyl-containing dimethylamine methyltransferase to remain on the adjacent region (Figure 22 and Table 18). Similar to the members belonging to the 7th order *Methanomassiliicoccales* methanogens, PylS in MSBL1 SAGs also lack the N-terminal extension compared with the full version of gene in *Methanosarcinaceae*. Furthermore, the annotation of the MSBL1 SAGs suggested that the MSBL1 Pyl system proteins likely contain in-frame TAG codons which incorporate Pyl (see also Figure 22). We also observed similar cases in Pyl system proteins in *Methermicoccus shengliensis*, *Methanococcoides methylutens*, multiple *Methanohalophilus* species, *Methanomassiliicoccus luminyensis*, and *Acetohalobium arabaticum*. However, we acknowledge that the protein product from these genes requires further experimental validation. The direct involvement of unusual amino acids in enzymes responsible for their biosynthesis has been previously reported in the biosynthesis of selenocysteine, the 21st unusual amino acid (Stock *et al.*, 2010).

1) MSBL1 SAG 382A20



2) MSBL1 SAG 382A03

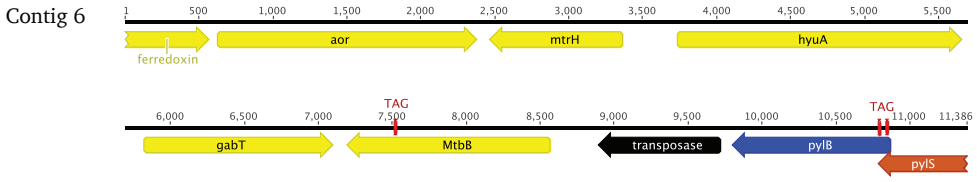


Figure 22. Schematic comparison of the putative Pyl synthesis and synthetase gene cluster region identified in the MSBL1 SAG 382A20 and SAG 382A03.

Table 18. Pyl systems in the MSBL1 SAGs.

SAG	Contig	Protein ID	Indigo annotation	Functional domain
382A20	44	AAA382A20_00858	Pyrrolysine--tRNA ligase protein	Pyrolysyl-tRNA synthetase, PylS
	44	AAA382A20_00857	Biotin synthase protein	Pyrrolysine biosynthesis radical SAM protein, PylB
	44	AAA382A20_00856	Phosphoribosylglycinamide formyltransferase 2 protein (PurT)	Pyrrolysine biosynthesis protein PylC
	44	AAA382A20_00855	Hypothetical protein	Pyrrolysine biosynthesis protein PylD
	202	AAA382A20_01503	tRNA-other	Pyl-tRNA
382A03	6	AAA382A03_00127	Pyrrolysine--tRNA ligase protein	Pyrolysyl-tRNA synthetase, PylS
	6	AAA382A03_00126	Biotin synthase protein	Pyrrolysine biosynthesis radical SAM protein, PylB

SAG 382A20 comprises the following protein coding genes: a complete gene repertoire for Pyl synthesis (*pylSBCD* with the exception of tRNA-^{Pyl}), trimethylamine-corrinoid protein co-methyltransferase (*mttB*) that utilizes Pyl and trimethylamine corrinoid protein (*mttC*) which is associated with the

former for trimethylamine-depended methyltransfer (Ferguson and Krzycki, 1997). The above six genes are classified on three co-transcribed operons: *pylS* and *pylB*, *pylD* and *pylC*, and *mtmB* and *mttC*. We identified basal promoter element TATA box sequences as listed in Table 19. The identified transcription factor B recognition elements (BRE) are shown in Figure 22. In SAG 382A20, PylD could have at least two potential translation versions predicted from its genome sequence. The default version was generated by the INDIGO annotation using GTG as translation initiator. The GTG-initiating version's PylD enzyme comprises 312 amino acids, of which the 5' has a fragment of amino acids extension longer by 32 to 40 amino acids when compared to all its homolog proteins in the NCBI non-redundant GenBank CDS database. This transcript containing PylD and PylC does not contain an obvious TATA box in its 5' UTR. In contrast, the other version of the PylD enzyme comprises 272 amino acids using ATG as translation initiator. In this case, two predicted TATA-boxes are observed in its upstream region suggesting that these MSBL1 genes are possibly under the archaeal regulatory machinery. Both versions are subject to confirmation or rejection by future experiments.

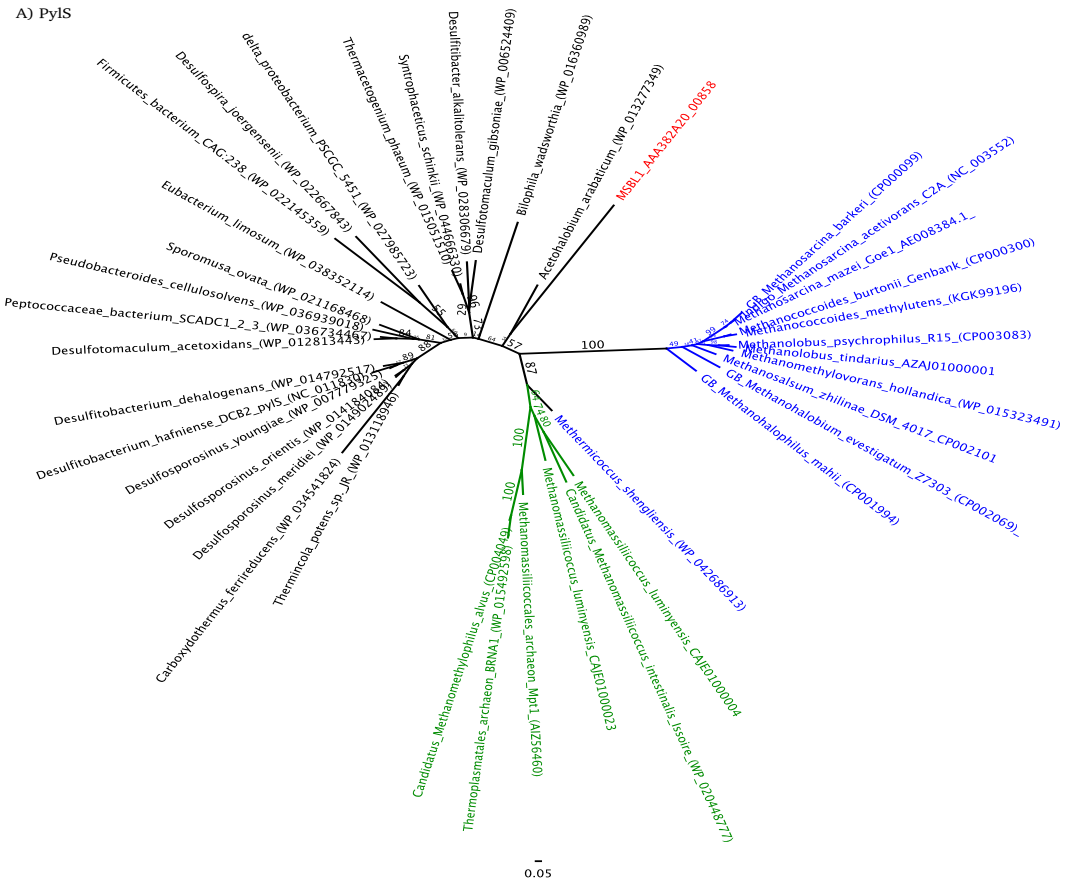
Table 19. Identified promoter sequences related to the Pyl systems in MSBL1 SAGs.

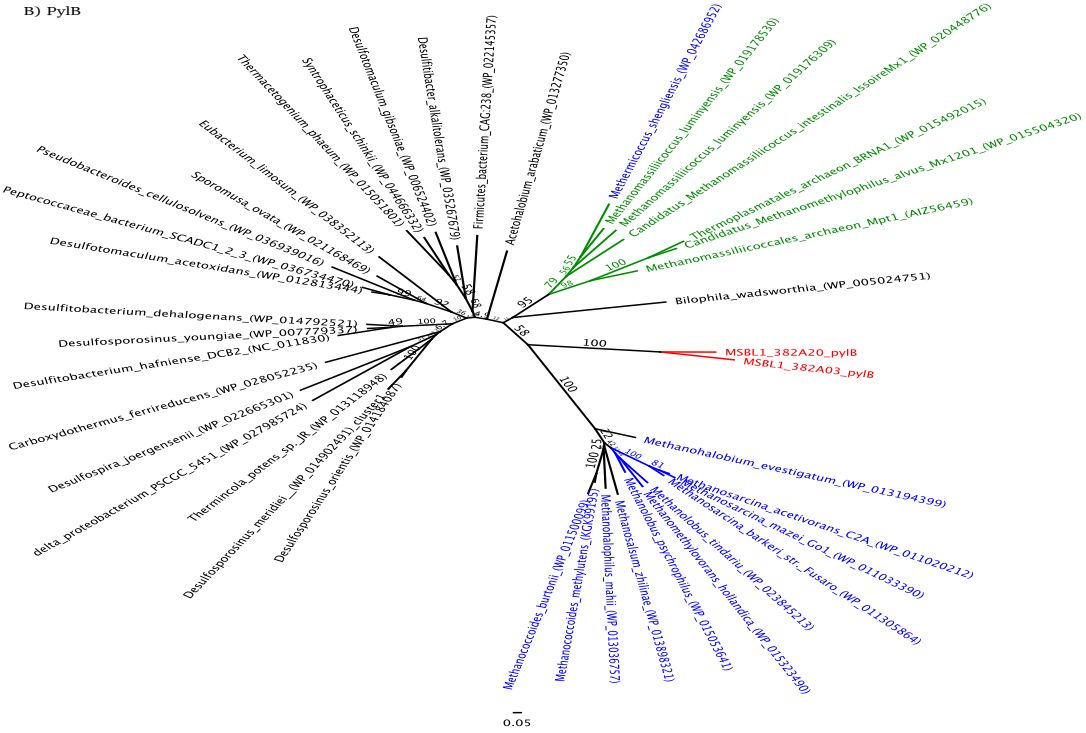
	MSBL1 SAG 382A20			
Enzyme	PylS	PylD GTG initiating	PylD ATG initiating	MtmB
Locus ID	82A20_00858	AAA382A20_00855	AAA382A20_00855	AAA382A20_00854
5' promoter type	TATA box	Not identified	TATA box	TATA box
Sequence	TAAAAG		TAAAAG	TAAATA
Nts until initiator	22~17 (ATG)		38~33 (ATG)	43~38 (ATG)
5' promoter type			TATA box	
Sequence			ATAAAA	
Nts until initiator			60~55 (ATG)	

	MSBL1 SAG 382A03
Enzyme	MtbB
Locus ID	AAA382A03_00124
5' promoter type	TATA box
Sequence	TATAAA
Nts until initiator	84~79 (ATG)
5' promoter type	
Sequence	
Nts until initiator	

Pyl system proteins in MSBL1 SAGs 382A20 and 382A03 have low amino acid sequence similarities (<46%) to their homologous proteins in cultivated or Candidatus species. Phylogenetic analysis of each of the predicted MSBL1 Pyl proteins in relationship with their homologs archived in the NCBI-non-redundant protein sequences (nr) database revealed the complex evolutionary history of Pyl system enzymes involving lateral gene transfers among taxa (see also Figure 23). Similar to previous findings (Herring *et al.*, 2007; Borrel, Gaci, *et al.*, 2014), the Pyl system proteins of members of the *Methanosarcinaceae* and members of *Methanomassiliicoccales* form two separate clusters, while the bacterial genes differ more markedly from their counterparts in *Methanosarcinaceae*. However, we identified that *Methermicoccus shengliensis*, a member of the order *Methanosarcinales*, has all four enzymes in its Pyl system closely related to the order *Methanomassiliicoccales,* thus the Pyl systems in *Methanosarcinales* are paraphyletic. Like the pylS genes in MSBL1 SAG 382A20 and genomes of *Methanomassiliicoccales*, *Methermicoccus shengliensis* also lacks the N-terminal of PylS. Four Pyl-system enzymes in MSBL1 are deep branching and do not have a consistent phylogenetic affiliation. It's highly likely that

these genes in MSBL1 have a mixed ancestry with homologous genes present in both bacteria and archaea. Similar cases are also observed in species of bacteria and many bacterial genera are paraphyletic albeit with insignificant support.




B) PylB
Methermicoccus_shengliensis_(WP_042686952)
Methanomassiliicoccus_luminyensis_(WP_019178530)
Methanomassiliicoccus_luminyensis_(WP_019176309)
Candidatus_Methanomassiliicoccus_intestinalis_IssoireMx1_(WP_020448776)
Thermoplasmatales_archaeon_BRNA1_(WP_015492015)
Candidatus_Methanomethylophilus_alvus_Mx1201_(WP_015504320)
Methanomassiliicoccales_archaeon_Mpt1_(AIZ56459)
Bilophila_wadsworthia_(WP_005024751)
MSBL1_382A20_pylB
MSBL1_382A03_pylB
Methanohalobium_evestigatum_(WP_013194399)
Methanosarcina_acetivorans_C2A_(WP_011020212)
Methanosarcina_mazei_Go1_(WP_011033390)
Methanosarcina_barkeri_str._Fusaro_(WP_011305864)
Methanolobus_tindarius_(WP_023845213)
Methanomethylovorans_hollandica_(WP_015323490)
Methanolobus_psychrophilus_(WP_015053641)
Methanosalsum_zhilinae_(WP_013898321)
Methanohalophilus_mahii_(WP_013036757)
Methanococcoides_methylutens_(KCK99519)
Methanococcoides_burtonii_(WP_011500098)
Desulfitobacterium_dehalogenans_(WP_014792521)
Desulfosporosinus_youngiae_(WP_007779327)
Desulfitobacterium_hafniense_DCB2_(NC_011830)
Carboxydothermus_ferrireducens_(WP_028052235)
Desulfospira_joergensenii_(WP_022665301)
delta_proteobacterium_PSCGC_5451_(WP_027985724)
Thermincola_potens_JR_(WP_013118948)
Desulfosporosinus_meridiei_(WP_014902491)_cluster1
Desulfosporosinus_orientis_(WP_014184068)
Peptococcaceae_bacterium_SCADC1_2_3_(WP_036739016)
Desulfotomaculum_acetoxidans_(WP_012813449)
Pseudobacteroides_cellulosolvens_(WP_036939016)
Sporomusa_ovata_(WP_021168469)
Eubacterium_limosum_(WP_038352113)
Thermacetogenium_phaeum_(WP_015051801)
Syntrophaceticus_schinkii_(WP_044866332)
Desulfotomaculum_gibsoniae_(WP_006524402)
Desulfitibacter_alkalitolerans_(WP_032567670)
Firmicutes_bacterium_CAC:238_(WP_022145357)
Acetohalobium_arabaticum_(WP_013277350)
0.05
C) PylC
AAA382A20_00355_pylC
Acetohalobium_arabaticum_(ADL11906)
Carboxydothermus_ferrireducens_(WP_028052234)
Desulfitobacterium_dehalogenans_(WP_014792522)
Desulfitobacterium_hafniense_DCB2_(WP_011462268)
Firmicutes_bacterium_CAC:238_(WP_022145358)
Methanococcoides_methylutens_(KCK99194)
Methanococcoides_burtonii_(WP_011500098)
Methanohalophilus_mahii_(WP_013036756)
Methanolobus_psychrophilus_(WP_015053642)
Methanolobus_tindarius_(WP_023845212)
Methanosarcina_acetivorans_(WP_011020211)
Methanosarcina_mazei_(WP_011033389)
Methanosarcina_barkeri_(WP_011305863)
Methanohalobium_evestigatum_(WP_013194398)
Methanosalsum_zhilinae_(WP_013898322)
Methanomethylovorans_hollandica_(WP_015323489)
Methanomassiliicoccus_luminyensis_(WP_026069110)
Methermicoccus_shengliensis_(WP_042686950)
Methanomassiliicoccus_luminyensis_(WP_019178531)
Methanomassiliicoccus_luminyensis_pylC_contig4
Candidatus_Methanomassiliicoccus_intestinalis_(WP_020448375)
Methanomassiliicoccales_archaeon_Mpt1_pylC
Thermoplasmatales_archaeon_BRNA1_(WP_015492007)
Candidatus_Methanomethylophilus_alvus_(WP_015505007)
Peptococcaceae_bacterium_SCADC1_2_3_(WP_036739017)
Desulfotomaculum_acetoxidans_(WP_012813448)
Pseudobacteroides_cellulosolvens_(WP_036939014)
Sporomusa_ovata_(WP_021168470)
Eubacterium_limosum_(WP_038352112)_
Desulfitibacter_alkalitolerans_(WP_028306682)
Desulfotomaculum_gibsoniae_(WP_006524403)
Syntrophaceticus_schinkii_(CEO90336)
Thermacetogenium_phaeum_(WP_015051512)
Desulfosporosinus_youngiae_(WP_007781549)
Desulfosporosinus_meridiei_(WP_014902490)
Desulfosporosinus_orientis_(WP_014184068)
Thermincola_potens_(WP_013118949)
delta_proteobacterium_PSCGC_5451_(WP_038240565)
Bilophila_wadsworthia_(WP_005024750)
Desulfospira_joergensenii_(WP_033798380)
0.05

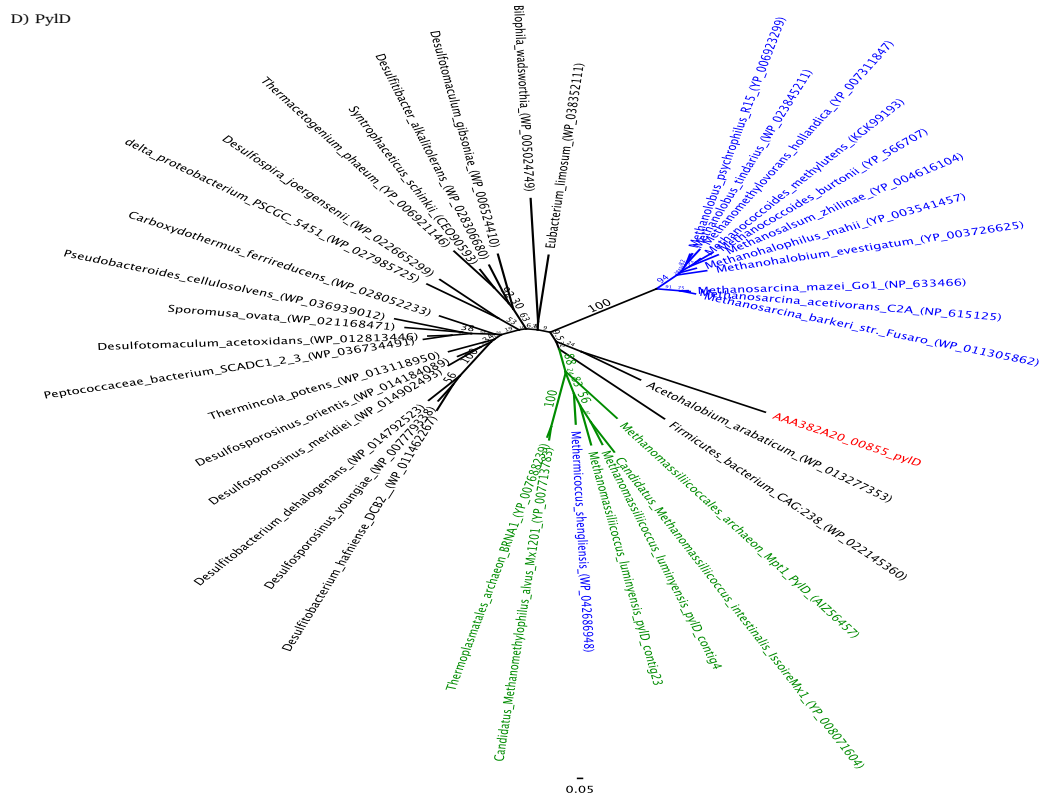


Figure 23. Phylogenetic inference of deduced peptide sequences of individual pylS/B/C/D gene. Sequences from MSBL1 SAGs are in red; sequences are of *Methanosarcinales* origin are shown in blue, sequences are of the 7th order methanogen *Methanomassiliicoccales* are shown in green; bacterial sequences are in black.

Unlike other Pyl-decoding archaea (methanogens) that are constitutively encoding UAG as Pyl regardless of carbon source (Heinemann *et al.*, 2009) and generally minimize the amber codon UAG frequency to only about 4% of the CDS (Prat *et al.*, 2012) (except for *Methanomassiliicoccus luminyensis* – 11.3% (Borrel *et al.*, 2014) and *Methermicoccus shengliensis* – 25%), MSBL1 SAGs contain about 23% UAGs in their CDS which is more similar to Pyl-decoding bacteria than to Pyl-decoding archaea (refer also to Table 20). Thus, we think it is likely that Pyl system genes in MSGL1 are regulated by substrates.

Table 20. Percentage of CDS with predicted UAG codon of the total CDS in the MSBL1 SAGs and various genomes of pyrrolysine coding microorganisms.

MSBL1	**UAG CDS, %**
382A03 (*pylS* containing)	17.0
382A20 (*pylS* containing)	21.1
382A13	20.9

Pyl-decoding methanogens within *Thermoplasmata*	**UAG CDS, %**
Thermoplasmatales archaeon BRNA1	2.0
Methanomassiliicoccales archaeon Mpt1	1.8
"*Ca.* Methanomethylophilus alvus", Mx1201*	2.8
"*Ca.* Methanomassiliicoccus intestinalis", Mx1-Issoire"*	5.3
*Methanomassiliicoccus luminyensis**	11.3

Pyl-decoding archaea belonging to *Methanosarcinales*	**UAG CDS, %**
Methermicoccus shengliensis	25
*Methanococcoides burtonii**	5.8
*Methanohalobium evestigatum**	4.1
*Methanohalophilus mahii**	4.1
*Methanosarcina acetivorans**	5.7

Pyl-decoding bacteria*	**UAG CDS, %**
Acetohalobium arabaticum	19.5
Desulfotomaculum acetoxidans	24.7
Desulfitobacterium dehalogenans	23.5

* Information obtained from Prat et al. (2012) and Borrel et al. (2014).

In MSBL1 SAG 382A20, enzymes required for the pathway to transfer the methyl group from various methylamines to methyl coenzyme M is complete (see also Table S1). In combination with Mwirichia *et al*'s analysis, MSBL1 SAGs have the metabolic capability for reversal of CO_2 reductive pathway. We propose that MSBL1 382A20 is able to oxidize the methyl group from various methyl donors to CO_2 to produce reducing equivalents.

Only two MSBL1 SAGs that possess Pyl decoding potential and this could be explained by;1) a technical issue in the multiple displacement amplification

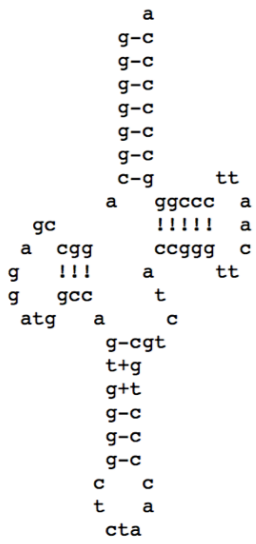
(MDA) technique that it might under-amplify certain genome regions from a single cell (as reviewed in (Huang *et al.*, 2015)). 2) The incompleteness of SAGs: all 32 MSBL1 SAGs reported by Mwirichia *et al.* are less than 60% complete (estimated). 3) Metabolic diversification and niche partition within the Candidate division MSBL1. The two MSGL1 SAGs (382A03 and 382A20) bearing a *pyl* gene cluster reported in this study are retrieved from Nereus Deep. SAG 382S20 is phylogenetically divergent from other MSBL1 SAGs analyzed (Mwirichia *et al*, accepted). It is possible that only certain populations of MSBL1 have the potential to decode pyrrolysine and utilize methylamines, while other MSBL1s do not.

4.4 Conclusions

In summary, with the advantage of single-amplified genomes and more sequenced genomes, there is an increase in the discovery of *pyl*-system genes. Here, we report members of MSBL1— one of the most abundant archaea in the deep-sea hypersaline brines of the Red Sea—to putatively decode pyrrolysine. We observed that the phylogeny of pyl-system genes do not necessarily follow the phylogeny of their source organisms. We also point out that the methylotrophic methanogen *Methermicoccus shengliensis* and putatively Pyl-decoding members of MSBL1 do not minimize their TAG codon frequency in the genome, indicating a potentially unusual mechanism for the regulation of Pyl synthesis and incorporation in these methanogenic archaea.

4.5 Supplementary materials

Figure S1. Predicted secondary structure of $tRNA^{Pyl}$ from MSBL1 SAG 382A20.




```
tRNA-Pyl(cta)
72 bases, %GC = 70.8
Sequence [136,207]
```

```
Primary sequence for tRNA-Pyl(cta)
1    .    10    .    20    .    30    .    40    .    50
ggggggcaggccgaggatggccagtggggctctaacccctgcgtctaccg
ggttcaattcccgggccccca
```

Table S1. Predicted enzymes related to methylamines (mono-, di-, and tri-), methanol and methylated thiols utilization in MSBL1 SAGs.

SAG ID	Trimethylamine corrinoid protein	Dimethylamine corrinoid protein	trimethylamine---corrinoid protein Co-methyltransferase	Methylated-thiol--corrinoid protein MtsB	Trimethylamine methyltransferase MttB	Dimethylamine methyltransferase MtbB2	Monomethylamine methyltransferase MtmB1	MethylcobamideCoM methyltransferase MtbA	methanol-5-hydroxybenzimidazolylcobamide co-methyltransferase
259A05			259A05_00835						
259B11	259B11_00805	259B11_00805							
259D18	259D18_00530		259D18_00414						
		259D18_00291	259D18_00411						
259E17									
		259E17_00564							
		259E19_01024							
259E19	259E19_00723	259E19_01201							259E19_01202
		259E19_00506							
259I09			259I09_00725						
259M10	259M10_00897		259M10_00896						
259O05	259O05_00631	259O05_00875							
261C02		261C02_00252							
261D19								261D19_00521	
								261D19_00522	
								261F19_00026	
261F19	261F19_00028							261F19_00027	
								261F19_00029	
261G05		261G05_00198							
261O19		261O19_00343							
					382A03_00049				
382A03		382A03_00640			382A03_00050	382A03_00124			
					382A03_00087	382A03_00123			
					382A03_00088				
					382A13_00477				
382A13			382A13_00401		382A13_00241				
			382A13_00402		382A13_00400				
			382A13_00533		382A13_00477				
								382A20_00951	
382A20	382A20_00952			382A20_00198			382A20_00854	382A20_00502	382A20_00199
	382A20_00852						382A20_00853	382A20_00265	
382C18				382C18_00485					382C18_00486
382K21					382K21_00265				

Table S2. Homologous proteins of methyl compound utilization-related enzymes in SAGs 382A20, 382A03, 382A13.

MSBL1 SAG 382A20

MtmB (AAA382A20_00854, AAA382A20_00853 and their intergenic region, manually modification with pyl incorporating), 473 aa	Max score	Total score	Query cover	E value	Ident	Accession
Monomethylamine:corrinoid methyltransferase [Candidatus Methanomassiliicoccus intestinalis Issoire-Mx1]	501	501	98%	7.00E-171	52%	AGY50174.1
Monomethylamine:corrinoid methyltransferase [Candidatus Methanomassiliicoccus intestinalis Issoire-Mx1]	501	501	98%	7.00E-171	53%	AGY50175.1
Monomethylamine methyltransferase MtmB [Thermoplasmatales archaeon BRNA1]	501	501	98%	9.00E-171	53%	YP_007688243.1
Monomethylamine:corrinoid methyltransferase; pyrrolysine-containing [Candidatus Methanomethylophilus alvus Mx1201]	488	488	98%	1.00E-165	53%	YP_007713787.1
monomethylamine:corrinoid methyltransferase MtmB [Methanomassiliicoccales archaeon Mpt1]	480	480	97%	1.00E-162	52%	AIZ56465.1
Monomethylamine:corrinoid methyltransferase; pyrrolysine-containing [Candidatus Methanomethylophilus alvus Mx1201]	479	479	98%	4.00E-162	50%	YP_007713786.1
Monomethylamine methyltransferase MtmB [Thermoplasmatales archaeon BRNA1]	477	477	97%	2.00E-161	50%	YP_007688242.1
monomethylamine:corrinoid methyltransferase [Acetohalobium arabaticum DSM 5501]	476	476	97%	9.00E-161	51%	ADL13224.1
monomethylamine methyltransferase MtmB1 [Methanomassiliicoccales archaeon Mpt1]	471	471	98%	5.00E-159	50%	AIZ56462.1
monomethylamine methyltransferase MtmB [Sporomusa ovata DSM 2662]	467	467	98%	1.00E-157	49%	EQB27844.1
Monomethylamine methyltransferase MtmB [Desulfotomaculum acetoxidans DSM 771]	462	462	98%	2.00E-155	50%	ACV60989.1
monomethylamine methyltransferase [Methanosarcina mazei Go1]	456	456	98%	4.00E-153	48%	NP_635359.2
RecName: Full=Monomethylamine methyltransferase MtmB; Short=MMA methyltransferase; Short=MMAMT [Methanosarcina mazei Go1]	456	456	98%	4.00E-153	48%	P58969.3
Monomethylamine methyltransferase MtmB [Methanosalsum zhilinae DSM 4017]	452	452	96%	9.00E-152	49%	YP_004616580.1
Monomethylamine methyltransferase MtmB [Methanosalsum zhilinae DSM 4017]	452	452	96%	1.00E-151	49%	YP_004615716.1
RecName: Full=Monomethylamine methyltransferase MtmB1; Short=MMA methyltransferase 1; Short=MMAMT 1 [Methanosarcina acetivorans C2A]	451	451	98%	4.00E-151	48%	P58865.3
monomethylamine methyltransferase [Methanosarcina acetivorans C2A]	451	451	98%	4.00E-151	48%	NP_615117.1
Monomethylamine methyltransferase MtmB [Methanolobus tindarius]	448	448	98%	3.00E-150	47%	WP_023845220.1
Monomethylamine methyltransferase MtmB [Methanolobus tindarius]	446	446	98%	4.00E-149	46%	WP_023845254.1
Monomethylamine methyltransferase MtmB [Methanolobus tindarius]	446	446	98%	4.00E-149	46%	WP_023845238.1
monomethylamine methyltransferase MtmB [Methanohalobium evestigatum Z-7303]	445	445	98%	6.00E-149	47%	YP_003726750.1

MttC (382A20_00852) 210 aa	Max score	Total score	Query cover	E value	Ident	Accession
dimethylamine methyltransferase [Methermicoccus shengliensis]	232	232	97%	3.00E-73	63%	WP_042686889.1
dimethylamine methyltransferase [Methanomassiliicoccus luminyensis]	221	221	99%	8.00E-69	55%	WP_026068678.1
putative cobalamin binding protein [Candidatus Methanomassiliicoccus intestinalis Issoire-Mx1]	210	210	99%	1.00E-64	51%	WP_020448778.1
hypothetical protein [Methanomassiliicoccus luminyensis]	197	197	99%	1.00E-59	52%	WP_019178518.1
monomethylamine corrinoid protein MtmC [Methanomassiliicoccales archaeon Mpt1]	192	192	100%	2.00E-57	48%	AIZ56466.1
putative cobalamin binding protein [Thermoplasmatales archaeon BRNA1]	189	189	99%	3.00E-56	46%	WP_015492601.1
dimethylamine methyltransferase [Methanomassiliicoccus luminyensis]	182	182	100%	2.00E-53	44%	WP_026068679.1
MULTISPECIES: Dimethylamine methyltransferase corrinoid protein [Candidatus Methanomethylophilus alvus Mx1201]	179	179	99%	1.00E-52	44%	WP_015505011.1
trimethylamine corrinoid protein [Candidatus Methanomassiliicoccus intestinalis Issoire-Mx1]	173	173	99%	4.00E-50	42%	WP_020448772.1
hypothetical protein [Methanomassiliicoccus luminyensis]	169	169	93%	1.00E-48	50%	WP_019178555.1
trimethylamine corrinoid protein [Candidatus Methanomethylophilus alvus Mx1201]	167	167	99%	9.00E-48	45%	WP_015504327.1
methyltransferase cognate corrinoid protein [Methanoculleus sp. CAG:1088]	167	167	99%	1.00E-47	45%	WP_022532207.1
methyltransferase [Acetohalobium arabaticum]	165	165	95%	7.00E-47	40%	WP_041667236.1
methyltransferase cognate corrinoid protein [Acetohalobium arabaticum DSM 5501]	165	165	95%	7.00E-47	40%	ADL11990.1
dimethylamine corrinoid protein 3 [Methermicoccus shengliensis]	163	163	100%	4.00E-46	43%	WP_042686907.1
corrinoid methyltransferase [Dictyoglomus thermophilum]	159	159	96%	7.00E-45	42%	WP_012547679.1

methyltransferase [Desulfosporosinus sp. BIC-A1/1_c9]	158	158	97%	2.00E-44	40%	WP_034145973.1
Trimethylamine corrinoid protein [Syntrophaceticus schinkii]	157	157	98%	4.00E-44	37%	CEO88859.1
trimethylamine corrinoid protein [Candidatus Methanomassiliicoccus intestinalis Issoire-Mx1]	157	157	98%	6.00E-44	40%	WP_020448782.1

MSBL1 SAG 382A03

MtbB (AAA382A03_00124, AAA382A03_00123 and their intergenic region, manually modification with pyl incorporating), 458 aa	**Max score**	**Total score**	**Query cover**	**E value**	**Ident**	**Accession**
Dimethylamine methyltransferase MtbB2[Methanosarcina acetivorans C2A]	443	443	99%	2.00E-148	50%	Q8TS72.3
dimethylamine methyltransferase [Methanosarcina acetivorans C2A]	443	443	99%	2.00E-148	50%	NP_615886.1
Dimethylamine methyltransferase MtbB1 [Methanosarcina mazei Go1]	434	434	99%	6.00E-145	49%	P58970.2
Dimethylamine methyltransferase MtbB2 [Methanosarcina barkeri str. Fusaro]	434	434	98%	1.00E-144	49%	P0C0W6.2
Dimethylamine methyltransferase MtbB3 [Methanosarcina acetivorans C2A]	431	431	99%	2.00E-143	48%	Q8TN68.3
dimethylamine methyltransferase [Methanosarcina acetivorans C2A]	431	431	99%	2.00E-143	48%	NP_617331.1
dimethylamine methyltransferase [Methanosalsum zhilinae DSM 4017]	430	430	99%	4.00E-143	48%	YP_004615691.1
Dimethylamine methyltransferase MtbB1 [Methanosarcina barkeri str. Fusaro]	429	429	98%	8.00E-143	49%	P0C0W5.2
Dimethylamine methyltransferase MtbB1 [Methanosarcina barkeri]	427	427	98%	5.00E-142	48%	O93661.3
Dimethylamine methyltransferase MtbB1[Methanosarcina acetivorans C2A]	422	422	98%	5.00E-140	48%	Q8TTA5.3
dimethylamine methyltransferase [Methanosarcina acetivorans C2A]	422	422	98%	5.00E-140	48%	NP_615496.1
dimethylamine:corrinoid methyltransferase [Methanolobus tindarius]	421	421	99%	2.00E-139	47%	WP_023845232.1
dimethylamine:corrinoid methyltransferase [Methanolobus tindarius]	421	421	99%	2.00E-139	47%	WP_023845252.1
dimethylamine methyltransferase [Methanohalobium evestigatum Z-7303]	420	420	99%	3.00E-139	49%	YP_003727089.1
Dimethylamine methyltransferase MtbB2 [Methanosarcina mazei Go1]	419	419	98%	6.00E-139	47%	P58971.2
dimethylamine methyltransferase [Methanohalobium evestigatum Z-7303]	419	419	99%	6.00E-139	48%	YP_003727571.1
Dimethylamine methyltransferase MtbB3 [Methanosarcina mazei Go1]	419	419	98%	1.00E-138	47%	P58972.2
dimethylamine:corrinoid methyltransferase [Methanomethylovorans hollandica DSM 15978]	418	418	99%	2.00E-138	47%	YP_007311861.1
dimethylamine:corrinoid methyltransferase [Methanomethylovorans hollandica DSM 15978]	418	418	99%	2.00E-138	47%	YP_007312794.1
dimethylamine methyltransferase MtbB1 [Sporomusa ovata DSM 2662]	411	411	98%	6.00E-136	45%	EQB27837.1
Dimethylamine:corrinoid methyltransferase; pyrrolysine-containing [Candidatus Methanomethylophilus alvus Mx1201]	411	411	99%	1.00E-135	44%	YP_007713794.1

MrtH (382A03_00120), 301 aa	**Max score**	**Total score**	**Query cover**	**E value**	**Ident**	**Accession**
methyltransferase [Methanothermus fervidus DSM 2088]	245	245	99%	2.00E-75	41%	WP_013413853.1
tetrahydromethanopterin S-methyltransferase subunit H [Desulfitobacterium hafniense]	242	242	99%	2.00E-74	40%	WP_018305631.1
tetrahydromethanopterin S-methyltransferase subunit H [Desulfitobacterium hafniense]	241	241	99%	4.00E-74	39%	WP_005816523.1
methyltransferase [Methanothermobacter marburgensis str. Marburg]	239	239	99%	3.00E-73	40%	WP_013296329.1
tetrahydromethanopterin S-methyltransferase subunit H [Methanothermobacter thermautotrophicus CaT2]	239	239	99%	4.00E-73	39%	BAM70298.1
methyltransferase [Methanothermobacter thermautotrophicus str. Delta H]	239	239	99%	4.00E-73	39%	WP_010876780.1
tetrahydromethanopterin S-methyltransferase subunit H [Archaeoglobus fulgidus DSM 4304]	236	236	92%	2.00E-72	45%	WP_010877523.1
N5-methyltetrahydromethanopterin:coenzyme M methyltransferase (mtr) {Methanobacterium thermoautotro [Archaeoglobus fulgidus DSM 4304]	236	236	92%	2.00E-72	45%	AAB91219.1
tetrahydromethanopterin S-methyltransferase subunit H [Thermincola potens]	236	236	99%	3.00E-72	39%	WP_013121766.1
MULTISPECIES: tetrahydromethanopterin S-methyltransferase subunit H [Desulfitobacterium]	234	234	99%	2.00E-71	38%	WP_014795441.1
tetrahydromethanopterin S-methyltransferase subunit H [Desulfosporosinus sp. BIC-A1/1_c9]	231	231	99%	3.00E-70	38%	WP_034141738.1
tetrahydromethanopterin S-methyltransferase subunit H [Desulfosporosinus meridiei]	231	231	99%	3.00E-70	38%	WP_014901418.1
Tetrahydromethanopterin S-methyltransferase subunit H [Methanobacterium sp. MB1]	230	230	99%	9.00E-70	39%	WP_023992793.1
methyltransferase [Methanobacterium sp. SWAN-1]	230	230	92%	9.00E-70	40%	WP_013826556.1
tetrahydromethanopterin S-methyltransferase subunit H [Desulfosporosinus orientis]	229	229	99%	1.00E-69	37%	WP_014182941.1
methyltransferase 2 [Sporomusa ovata]	229	229	99%	3.00E-69	37%	WP_021169914.1

MttB (382A03_00049, 382A03_00050 and their intergenic region, manually modification with pyl incorporating), 488 aa	**Max score**	**Total score**	**Query cover**	**E value**	**Ident**	**Accession**
trimethylamine methyltransferase [Methanohalobium evestigatum]	340	340	98%	3.00E-107	40%	WP_013195346.1
Trimethylamine methyltransferase MttB [Methanosarcina thermophila]	332	332	94%	3.00E-104	41%	Q9P995.4
trimethylamine corrinoid protein methyltransferase MttB [Methanosarcina thermophila TM-1]	332	332	94%	3.00E-104	41%	AAD38789.1
Trimethylamine methyltransferase MttB [Methanosarcina barkeri str. Fusaro]	330	330	96%	1.00E-103	40%	P0C0W7.3

trimethylamine:corrinoid methyltransferase ; pyrrolysine-containing [Candidatus Methanomassiliicoccus intestinalis Issoire-Mx1]	330	330	98%	2.00E-103	40%	AGY50177.1
trimethylamine:corrinoid methyltransferase [Methanolobus tindarius]	330	330	98%	2.00E-103	39%	WP_023845246.1
Trimethylamine methyltransferase MttB [Methanosarcina barkeri]	329	329	96%	3.00E-103	40%	O93658.4
Trimethylamine methyltransferase MttB1 [Methanosarcina acetivorans C2A]	324	324	96%	5.00E-101	39%	Q8TTA9.4
trimethylamine methyltransferase [Methanosarcina acetivorans]	324	324	96%	5.00E-101	40%	WP_011020577.1

MttB (382A03_00087, 382A03_00088 and their intergenic region, manually modification with pyl incorporating), 479 aa	**Max score**	**Total score**	**Query cover**	**E value**	**Ident**	**Accession**
Trimethylamine methyltransferase MttB [Methanosarcina barkeri str. Fusaro]	343	343	98%	2.00E-108	41%	P0C0W7.3
Trimethylamine methyltransferase MttB [Methanosarcina barkeri]	342	342	98%	4.00E-108	41%	O93658.4
Trimethylamine methyltransferase MttB [Methanosarcina thermophila]	337	337	96%	2.00E-106	41%	Q9P995.4
trimethylamine corrinoid protein methyltransferase MttB [Methanosarcina thermophila TM-1]	337	337	96%	2.00E-106	41%	AAD38789.1
Trimethylamine methyltransferase MttB1 [Methanosarcina acetivorans C2A]	335	335	98%	2.00E-105	40%	Q8TTA9.4
trimethylamine methyltransferase [Methanosarcina acetivorans C2A]	335	335	98%	2.00E-105	40%	NP_615492.1
trimethylamine:corrinoid methyltransferase [Methanolobus tindarius]	334	334	98%	4.00E-105	40%	WP_023845246.1

MSBL1 SAG 382A13

MttB (AAA382A13_00401, AAA382A13_00400 and their intergenic region), 475 aa	**Max score**	**Total score**	**Query cover**	**E value**	**Ident**	**Accession**
Trimethylamine methyltransferase MttB [Methanosarcina barkeri str. Fusaro]	338	338	98%	7.00E-107	40%	P0C0W7.3
Trimethylamine methyltransferase MttB [Desulfitobacterium hafniense DCB-2]	335	335	97%	8.00E-106	38%	Q18TV3.3
trimethylamine methyltransferase [Methanohalobium evestigatum]	335	335	99%	1.00E-105	39%	WP_013195346.1
Trimethylamine methyltransferase MttB [Methanosarcina barkeri]	335	335	98%	2.00E-105	40%	O93658.4
trimethylamine:corrinoid methyltransferase [Methanolobus tindarius]	334	334	98%	3.00E-105	40%	WP_023845246.1
Trimethylamine methyltransferase MttB1 [Methanosarcina acetivorans C2A]	329	329	98%	2.00E-103	39%	Q8TTA9.4
trimethylamine methyltransferase [Methanosarcina acetivorans]	329	329	98%	2.00E-103	39%	WP_011020577.1
Trimethylamine methyltransferase MttB2 [Methanosarcina acetivorans C2A]	327	327	98%	3.00E-102	40%	Q8TS73.3
trimethylamine methyltransferase [Methanosarcina acetivorans]	327	327	98%	3.00E-102	40%	WP_011020970.1

4.6 References

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. (1990). Basic local alignment search tool. *Journal of Molecular Biology* **215**: 403–410.

Borrel G, Gaci N, Peyret P, O'Toole PW, Gribaldo S, Brugere J-F. (2014). Unique characteristics of the pyrrolysine system in the 7th order of methanogens: implications for the evolution of a genetic code expansion cassette. *Archaea* **2014**: 1–11.

Borrel G, Parisot N, Harris HMB, Peyretaillade E, Gaci N, Tottey W, *et al.* (2014). Comparative genomics highlights the unique biology of Methanomassiliicoccales, a Thermoplasmatales-related seventh order of methanogenic archaea that encodes pyrrolysine. *BMC Genomics* **15**: 679.

Collingridge PW, Kelly S. (2012). MergeAlign: improving multiple sequence alignment performance by dynamic reconstruction of consensus multiple sequence alignments. *BMC Bioinf* **13**: 117.

Ferguson DJ, Krzycki JA. (1997). Reconstitution of trimethylamine-dependent coenzyme M methylation with the trimethylamine corrinoid protein and the isozymes of methyltransferase II from Methanosarcina barkeri. **179**: 846–852.

Gaston MA, Jiang R, Krzycki JA. (2011). Functional context, biosynthesis, and genetic encoding of pyrrolysine. *Current Opinion in Microbiology* **14**: 342–349.

Guan Y, Hikmawan T, Antunes A, Ngugi D, Stingl U. (2015). Diversity of methanogens and sulfate-reducing bacteria in the interfaces of five deep-sea anoxic brines of the Red Sea. *Research in Microbiology*. doi:10.1016/j.resmic.2015.07.002.

Guindon S, Dufayard J-F, Lefort V, Anisimova M, Hordijk W, Gascuel O. (2010). New

algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Systematic Biology* **59**: 307–321.

Heinemann IU, O'Donoghue P, Madinger C, Benner J, Randau L, Noren CJ, *et al.* (2009). The appearance of pyrrolysine in tRNAHis guanylyltransferase by neutral evolution. *Proceedings of the National Academy of Sciences* **106**: 21103–21108.

Herring S, Ambrogelly A, Polycarpo CR, Söll D. (2007). Recognition of pyrrolysine tRNA by the Desulfitobacterium hafniense pyrrolysyl-tRNA synthetase. *Nucleic Acids Res* **35**: 1270–1278.

Huang L, Ma F, Chapman A, Lu S, Xie XS. (2015). Single-cell whole-genome amplification and sequencing: methodology and applications. *Annu Rev Genomics Hum Genet* **16**: 79–102.

Katoh K, Kuma K-I, Toh H, Miyata T. (2005). MAFFT version 5: improvement in accuracy of multiple sequence alignment. *Nucleic Acids Res* **33**: 511–518.

Lang K, Schuldes J, Klingl A, Poehlein A, Daniel R, Brune A. (2015). New mode of energy metabolism in the seventh order of methanogens as revealed by comparative genome analysis of 'Candidatus Methanoplasma termitum' Elliot, MA (ed). *Appl Environ Microbiol* **81**: 1338–1352.

Laslett D, Canback B. (2004). ARAGORN, a program to detect tRNA genes and tmRNA genes in nucleotide sequences. *Nucleic Acids Res* **32**: 11–16.

Le SQ, Gascuel O. (2008). An improved general amino acid replacement matrix. *Mol Biol Evol* **25**: 1307–1320.

Prat L, Heinemann IU, Aerni HR, Rinehart J, O'Donoghue P, Söll D. (2012). Carbon source-dependent expansion of the genetic code in bacteria. *Proceedings of the National Academy of Sciences* **109**: 21070–21075.

Srinivasan G, James CM, Krzycki JA. (2002). Pyrrolysine encoded by UAG in Archaea: charging of a UAG-decoding specialized tRNA. *Science* **296**: 1459–1462.

Stock T, Selzer M, Rother M. (2010). In vivo requirement of selenophosphate for selenoprotein synthesis in archaea. *Molecular Microbiology* **75**: 149–160.

van der Wielen PWJJ. (2005). The enigma of prokaryotic life in deep hypersaline anoxic basins. *Science* **307**: 121–123.

CONCLUDING REMARKS AND FUTURE DIRECTIONS

This dissertation presents a comprehensive study of archaea in the deep-sea hypersaline habitats of the Red Sea, including their composition, the first report of a successful cultivation attempt for methanogens from the Red Sea followed by in-depth genomic explorations as well as a transcriptomic study. In order to elucidate the diversity, function, and adaptation of archaea, this dissertation further describes the methodology that enables the aforementioned contributions to the field of deep-sea microbiology, namely a combination of microbial community assessment as well as comparative genomics and transcriptomics approaches.

The archaeal community in the sampled deep-sea hypersaline habitats of the Red Sea is enriched with *Thaumarchaeota* and *Thermoplasmata* (primarily *Candidatus* Mediterranean Sea Brine Lakes group 1(MSBL1)). Typical methanogenic members are ubiquitous but in very low abundance and there was a considerable diversity of uncultivated and undescribed archaea. We subsequently conducted an intensive anaerobic cultivation investigation aiming to obtain novel representatives, especially methanogens, from the Red Sea brines. Two pioneer halophilic methanogens have been isolated from Kebrit Deep and Erba Deep, representing novel species affiliated with the genera *Methanohalobium* and *Methanohalophilus*, respectively. Comparative genomic analysis revealed their predicted capacities in osmo-adaptation, unique virus defense systems, distinctive membrane signatures, various

transport systems and DNA repair mechanisms that possibly enable them to thrive in the polyextreme environment. Multiple cellular processes in deep-sea methanogens are involved in coping with unfavorable salinities at the transcriptome level, primarily featuring, on the one hand, up-regulating genes for inorganic compound transport, amino acid metabolism ribosome proteins, and on the other hand, down-regulating genes for methanogenesis and membrane-bound NADH-quinone oxidoreductase, V-type ATP synthase and electron transport Rnf complexes.

Single-amplified genome (SAG) is a powerful approach to investigate the genome of the vast majority of uncultivated microorganisms. The uncultivated Mediterranean Sea Brine Lakes group 1(MSBL1) has been suggested as putative methanogens in numerous studies. We found that members of MSBL1 are putative methylamine utilizing and pyrrolysine decoding microbes. This finding has expanded the hypothetical substrate pool and the coding potential of this abundant and not yet cultivated lineage.

In summary, this dissertation has reinforced the observation of certain archaeal lineages that are frequently associated with the deep-sea brine habitat and has furthered our understanding of archaeal communities in the deep-sea brines of the Red Sea. Also it has established the first anaerobic culture collection from this environment and provided the first insights in how deep-sea methanogens could adapt to the polyextreme brine habitat. Understanding how methanogens cope with harsh habitats in the Red Sea will also contribute to our understanding of the evolution of archaea in extreme environments.

Future perspectives should embrace *in situ* enzyme activity measurement and single-cell derived genomes with greater coverage will be furthermore desired to provide a better picture of the metabolic capacity and ecological function of the majority of uncultured archaea in the deep-sea brines of the Red Sea. Future enrichment and cultivation attempts could be focused on cultivating the *Methanomassiliicoccales* associated methanogens as well as lineages that have no cultured representatives from the deep-sea brines.