Addis Ababa University
School of Graduate Studies
Institute of Biotechnology

Metagenomic Analysis of Microorganisms in Hyper Saline Aquatic and Acidic Habitats of Danakil Depression of the Afar Region of Ethiopia and Their Biotechnological Potential

PhD Dissertation

Lulit Tilahun Wolde

December, 2021

Addis Ababa, Ethiopia

Metagenomic Analysis of Microorganisms in Hyper Saline Aquatic and Acidic Habitats of Danakil Depression of the Afar Region of Ethiopia and Their Biotechnological Potential

Lulit Tilahun Wolde

A Thesis Submitted to the Institute of Biotechnology, School of Graduate Studies of the Addis Ababa University in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy in Biotechnology

Supervisors: Addis Simachew (PhD)
Asfawossen Asrat (Prof)
Gary M. Wessel (PhD)

Declaration

I, Lulit Tilahun Wolde, hereby declare that this dissertation and its entirety is my bona fide work and has been submitted in partial fulfillment of the requirements for a PhD degree at Addis Ababa University. This dissertation or its part has not been submitted to other University anywhere for the award of any degree. All sources of materials used for the dissertation have been duly acknowledged.

December, 2021

Thesis approved by supervisors for submission:

	Name	Signature	Date
1.	Addis Simachew (PhD)		
2.	Asfawossen Asrat (Prof)		
3.	Gary M. Wessel (PhD)		
4.	Diriba Muleta (PhD)		
5.	Awoke Guadie (PhD)		

Acknowledgments

Several people and Institutions have contributed to the completion of this Dissertation, and I sincerely acknowledge all of them for their understanding and timely support. My heart felt gratitude and cordial appreciation goes to my advisors, Dr. Addis Simachew from Institute of Biotechnology (IoB), Addis Ababa University, Professor Asfawossen Asrat from School of Earth Sciences, Addis Ababa University and Dr Gary M. Wessel from Department of Molecular and Biomedical, Brown University, Providence USA. They taught me not only research assistantship but also the strength to work independently and how to approach people who need support in the research endeavor.

My special acknowledgment is extended to all my advisers who believed in me and allowed me to follow my ambition. Particularly my deepest gratitude is to Dr Addis Simachew and Professor Asfawossen Asrat who followed me to the one of the harshest environments in the world and helped me to collect the important samples. I have special thanks to Dr. Gary M. Wessel who hosted me to work most of my experiments in his laboratory as well as took me in his home as part of his family and supported my entire 3 month stay in Providence USA. Dr. Wessel and his team at PRIMO laboratory have closely helped me during the experiment and molecular data analysis. My deepest gratitude to my three advisers for their unreserved and punctually provided scientific advisory during the course of different experiments and invaluable comments during manuscript preparations and Thesis write-ups.

The IoB directors Dr. Kassahun Tesfaye (former) and Dr. Tesfaye Sisay as well as the deputy director Dr. Addis Simachew have unreservedly helped me in different administrative issues, thank you. I also extend my acknowledgment to Dr. Sebastian Leiden and his team at Max Plank institute, Germany for aiding me in taking important imagery of biological samples of Dallol and Dr. Teklehaimanot H/Selassie for his assistance to facilitate accommodation and working place at Yara Dallol.

My dearest gratitude and hats off to my family, PhD fellow classmates, colleagues, close relatives and friends for their encouragements, scientific discussions and the good times we had over the past several years. Above all, my praises and thanks are reserved to my shepherd, the Almighty God for guiding and helping me all the way through my life particularly from the beginning to the end of this Thesis work.

All happened with the Will of God!

Table of Contents

Figure Index

Index of Tables

List of Acronyms and abbreviations

16S rRNA	16S ribosomal RNA
BLASTx	Basic Local Alignment Translated Search Tool
COD	Chemical Oxygen Demand
Contigs	Contiguous sequences
CTAB	Cetyltrimethyl Ammonium Bromide
Cys	Cysteine
DIAMOND	Double Index Alignment of Next Generation Data
DNA	Deoxyribonucleic Acid
eDNA	environmental DNA
EARS	East African Rift System
EMV	Extra Membrane Vesicles
homo-Cys	homo Cysteine
HsAP	Hypersaline Acid Ponds
HsAqS	Hypersaline Aquatic Systems
ITS	Internal Transcribed Spacer
KEGG	Kyoto Encyclopedia of Genes and Genomes
LCA	Lower Common Ancestor
MDS	Multidimensional Scaling
MEGAN6 CE	MEtaGenome Analyzer 6 Community Edition
MetaPhlAn	Metagenomic Phylogenetic Analysis

mNGS	metagenomic Next Generation Sequence
NCBI-nr	NCBI non-redundant protein database
OA-ICOS	Off-Axis Integrated Cavity Output Spectroscopy
ORFs	Open Reading Frames
OTU	Operational Taxonomic Unit
PcoA	Principal coordinate Analysis
PCR	Polymerase chain reaction
PrIMO	Providence Institute of Molecular Oogenesis
SDS	Sodium Dodecyl Sulfate
TE buffer	Tris-EDTA buffer
TEM	Transmission Electron Microscopy
TP	Total Phosphorous
UTM	Universal Transverse Mercator coordinate system
WMS	Whole-Metagenome Shotgun

Metagenomic analysis of microorganisms in hyper saline aquatic and acidic habitats of Danakil Depression of the Afar Region of Ethiopia and their Biotechnological Potential

Lulit Tilahun Wolde

PhD Dissertation

Addis Ababa University, 2021

Abstract

The Danakil Depression is one of the hottest place on Earth and is considered as an exceptional place where different extreme environments are found. Five locations in the Danakil Depression were studied for the presence of extremophiles and their biotechnological potentials. The sample sites for this study are the Hyper-saline Acid Ponds [HsAP] (Dallol [DAL], Black Water [BLA], Gaet'ale [GAL]) as well as the Hyper-saline Aquatic Systems [HsAqS] (coastal parts of Lake Asale [LA, LADP5, LADP6], Mud'ara pond [MUP]). All the sites were located South to Mount Dallol at less than -100 m below sea level. In our investigation, we tested for the presence of DNA-based life in the studied sites using Whole Metagenome Shotgun (WMS) DNA sequencing approaches. Bioinformatic tools that adopt techniques of assembly free (Bowtie, MetaPhlAn2) and assembly (MetaSpades, DIAMOND and MEGAN 6-CE) were used to identify the microbial communities and predict the major nutrient metabolic

networks. Accordingly, the HsAqS were found to be mainly dominated by archaea Operational Taxonomic Units (OTUs), while the HsAP were dominated by bacteria OTUs. More than 95% of the total OTUs of HsAqS was grouped under phylum Euryarchaea, while more than 90% of the OTUs of HsAP was classified in the phyla Proteobacteria, Actinobacteria and Firmicutus. *Halonotius* sp *J07HN6* and unclassified *Salinibacter ruber* were the two abundant archaeal and bacterial OTUs in the HsAqS, respectively. On the other hand, unclassified strain of *Paraburkholderia fungorum* was profiled as the most abundant OTU for DAL and BLA, while unclassified *Escherichia* and unclassified *Burkholderia* were top two abundant OTUs for GAL. Insilico prediction of gene functions and metabolic pathways using SEED and KEGG indicated presence of carbon fixation, sulfate oxidation, assimilative nitrogen and sulfate reduction reactions in DAL and BLA samples. Carbon fixation was also predicted in the HsAqS while methanogenesis was exclusively predicted in GAL. Based on the investigation of the sources of the genes of interest, it is clear that syntrophic metabolism is the main strategy adapted to utilize the nutrients in the studied sampling sites. Moreover, the predicted genes from DAL's and BLA's nitrogen metabolism helped to discover new protein sequences of cysteine desulfurase and [4Fe-4S] ferredoxin with less than 80% similarity to the homologous reference protein sequence. Images of DAL extremophiles obtained from Transmission Electron Microscopy (TEM) showed high production of nano sized Extracellular Membrane Vesicles (EMV). Further investigation on resistance of DAL's extremophiles to phenolic compounds under lysing conditions indirectly signified their

bio-remediation capability. In general, this preliminary investigation indicates great biotechnological and bioremediation potentials of extremophiles of Danakil Depression.

Key words; Danakil Depression, Extremophiles, Metagenomis, Insilico, Extreme environment

Chapter One

1. Introduction

1.1. Background of the study

The repertoire of habitats for biological occupancy are vast and varied. Our own environmental requirements for proliferation are relatively narrow, and they are greatly assisted by clothing and accessories. The unseen world of microorganisms in their natural environment with all the interactions within and between species has been the core interest of microbiologists for several decades. It is essential to know the potential of specific organism through its interactions to the closest and/or very distant relatives. This knowledge help to understand the exchange of resources and metabolic networks among different organisms to maintain the natural balance of the habitat. Furthermore, in order to set up and understand the limit of life, it is crucial to study life in extreme environments in the sense of both physical (geographic) as well as the capacity of life to hold out and adapt to change (Rothschild and Mancinelli, 2001; Rampelotto, 2013; Merino *et al.*, 2019). Hence, studying the taxonomical compositions of microbial communities and the possible “micro” social interactions in the environment is crucial since microorganisms are responsible for variety of biological activities occurring on earth.

Prokaryotes generally dominate the Earth's biosphere in comparison to other organisms, and their estimated total number inhabiting both normal and extreme habitats is about 4–6 × 10^{30} cell (Whitman *et al.,* 1998; Bunge *et al.,* 2014). In the past century, morphological and biochemical characterizations of microorganisms through culturing were used to study different types of isolates and shed light on the structure of microbial populations from various habitats. However, only a handful of prokaryotes (0.1% to 1%) were found to be cultivable. This is because it is not possible to replicate exact culture conditions of the sample environment and successfully grow all microorganisms in a sample. Moreover, evolution dictates rapid propagation of microbial cells primarily in their familiar natural environments otherwise they enter into dormant stage to lessen vulnerability to toxicity (Lewis 2009, p.182; Selvarajan *et al.*, 2014).

Nowadays, the advancement in molecular and computational technologies enable scientists to infer about the presence of organisms and their function in different habitats such as human guts and hot springs (Schmidt *et al.*, 1991; Narasingarao *et al.,* 2012; Luk *et al.,* 2015; Tan *et al.,* 2015; Ibarbalz *et al.,* 2016). The development of culture independent tools such as metagenomics, metatranscriptomics and metaproteomics helps to collect and study the largely untapped microbial resources (Thomas *et al.,* 2012). Especially in extreme conditions, most organisms instead rely on creative gene sets to assist survival and/or proliferation in the environments (Reed *et al.*, 2013; Coker, 2016; Orellana *et al.*, 2018; Donato *et al.*, 2019).

These machine based technologies may not have the capability to replace culture-based methods in microbial studies, but can be supplemental tools to cover the molecular characteristics. Using metagenomics which involve direct environmental sample, one can avoid certain problems inherent in culturing such as contamination, population congestion and so on. Thus, many extreme environments have been studied around the world for the presence of organisms that thrive in extreme conditions.

In general, metagenomics studies have shed light on novel aspects of biology, such as the possible connections between the human gut microbiome and diseases as diverse as diabetes, depression and so on (Lindgreen *et al.*, 2016). Most metagenome-based studies are centered on assessing the genomic composition and diversity within and across microbial communities. Such analyses can be done either by targeting 16S rRNA genes sequencing in bacteria and 18S rRNA genes for eukaryotes or by targeting the whole genome using WMS sequencing (Schmidt *et al.*, 1991; Handelsman *et al.*, 1998; Segata *et al.*, 2013). Instead of just studying one particular marker gene such as 16S rDNA, shotgun metagenomics sequencing involves sequencing the total microbial DNA of a sample and can be used to infer the relative abundance of every microbial gene and quantify specific metabolic pathways to predict the potential functionality of the entire community. One of the many projects of metagenomics was important for recognizing the vast functional gene diversity in the microbial worlds (Handelsman *et al.*, 1998).

By comparing shotgun metagenomics data across samples, it is possible to study larger scale issues such as human-microbe association, ecology and biogeography (Liu *et al.*, 2012^{a}; Narasingarao *et al.*, 2012; Stivaletta *et al.*, 2012; Tan *et al.*, 2015; Gibson *et al.*, 2016). Not only metagenomics, but other “meta” omics studies such as the metatranscriptomics, metaproteomics and metabolomics help to study the complex ecosystems microbes inhabit by revealing the genes, transcripts, proteins and the metabolites correspondingly for studying the systems-level microbial interactions and biochemical functions (Segata *et al.*, 2013).

To date, there is a global need to transit from polluting solvent based chemical reactions to a more ‘green’ and environmentally friendly alternative approaches in the industry sectors such as pharmaceutical, food, detergents, textiles, leather, pulp and paper (Sheldon, 2017; Chapman *et al.*, 2018). Biocatalysts from both eukaryote as well as prokaryote sources have been used in different industries for more than a century (Matsumoto *et al.*, 2001; Röhm, 1915 cited in Quax, 2006; Guan *et al.*, 2014; Sonia *et al.*, 2016). Recent studies have focused on prokaryotes from extreme environments to solve the recurring difficulty of stabilizing enzymes at high temperatures or in turbulent flow regimes, as well as in potentially toxic solvents (Chakravorty *et al.*, 2017; Chapman *et al.* 2018).

Biocatalysts such as enzymes have shown many unique advantages compare to the non-biological counterparts and are highly efficient for industrial scale catalysis. As enzymes are exceedingly more product selective, functional in milder reaction conditions and mitigate minimum environmental and physiological toxicity, their advantages can be translated into reduced operating costs when they were effectively employed as biocatalysts in chemical processes in different industries (Sheldon, 2017; Chapman *et al.*, 2018).

In this study, metagenomics was applied to characterize the diversity of microorganisms in the hyper saline and acidic hydrogeothermal as well as hyper saline aquatic environments of Danakil Depression. Furthermore, through mapping of metabolic pathways by identifying important enzymatic enzyme coding genes from the metagenome data, major biogeochemical processes were predicted in the studied different extreme environments.

1.2. Objectives of the study

1.2.1 General objective

The overall objective of this study was to determine the diversity of microbial communities in lake and hot springs found in the Danakil Depression of Afar region of Ethiopia using metagenomics.

1.2.2. Specific objectives

The specific objectives of the current study were:

- To examine the microbial diversity in Dallol acid Pond, Gaet’ale acid Pond, Black Water acid Pond, Lake As’ale and Mud’ara Pond using metagenomic approach.
- To predict and determine the nutrient metabolism pathways of microbial communities in Dallol, Gaet’ale and Black Water acid Ponds as well as Lake As’ale and Mud’ara Pond using insilico approaches with special emphasis on nitrogen cycle in Dallol and Black water acid ponds.
- To predict the biotechnological potential applications of the extremophiles inhabiting Dallol pond.

Chapter Two

2. Literature Review

2.1. Microbiome and extreme environments

Microbiome is defined as the totality of microbiota, the interactions of their structural elements (proteins, lipids, polysacharides and nucleic acids) and metabolites with each other and surrounding environment (Fig. 1) (Berg *et al.*, 2020). Numerous and different extreme microbiomes exist on Earth as they are mainly characterized by the types of inhabiting microbes (microbiota), that are directly affected by the extreme environmental conditions (Bragina *et al.,* 2014; Berg *et al.,* 2020; Rodriguez and Durán 2020; Berg *et al.*, 2021).

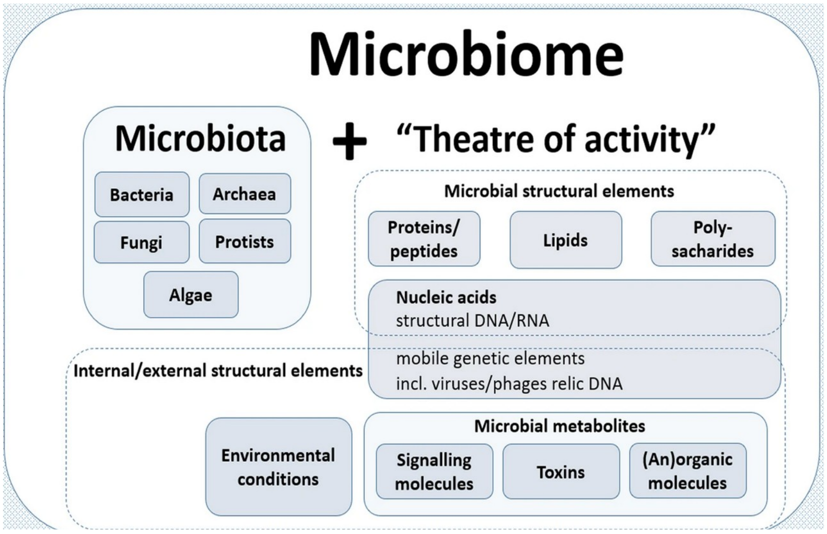


Fig. 1: Components of Microbiome.

The interactions between microbiota and their biological elements with the environment and/or among themselves. From: Microbiome definition re-visited: old concepts and new challenges (Berg et al., 2020)

Life in an extreme environment is very challenging from a human perspective. Almost all extreme environments are not suitable habitats for human beings or large animals except for a few that live in the Arctic (Rothschild and Mancinelli, 2001). Yet, some living organisms are found to be successfully dwelling in such environments either by thriving or just by tolerating it (Stout *et al.,* 2009; Jimenez *et al.,* 2012; Li *et al.,* 2014; Jimenez and Zambrano 2015). Sometimes extremophiles such as *Sulfolobus acidocaldarius* (an archaea that live in a pH 3 and 80 ^{0}C) can flourish in multiple extreme conditions and are referred to as polyextremophiles (Rothschild and Mancinelli, 2001).

Various survival strategies have been developed by extremophiles to overcome extreme environmental conditions. Thermophiles and pschrophiles modified their enzymes to be “tough” and “efficient” respectively (Hollien and Marqusee, 1999; Reed *et al.,* 2013; Wang *et al.,* 2015; Bala and Singh, 2018; Bhatia *et al.,* 2021). “Tough” enzymes of thermophiles have increased maximal stability & decreased heat capacity of unfolding (Δcp) while “Efficient” enzymes of pschrophiles have improved flexibility on the active site (Hollien and Marqusee, 1999; Feller, 2010; Feller, 2013; Reed *et al.,* 2013; Bhatia *et al.,* 2021). On the contrary, acidophiles, alkaliphiles and halophiles keep “normal” cell’s interior by excreting protective substances and/or enzymes to the exterior (Baker-Austin and Dopson, 2007; Reed *et al.,* 2013; Harding *et al.,* 2017; Mamo, 2020).

Extreme environments may harbor organisms from different taxonomical groups which are found in all kingdoms of life (Rothschild and Mancinelli, 2001). Many terrestrial hot springs found in different parts of the world had been well examined for their extremophile diversity using metagenomics, usually the 16 S rRNA gene analyses (Bartle, 2007; Jimenez *et al.,* 2012; Jimenez and Zambrano, 2015). Quantities and varieties of the peculiar organisms inhabiting extreme environments, such as the Dead sea lake and hot springs of Yellowstone are known to be significantly affected by physico-chemical factors such as salinity, pH, temperature and different ionic contents (Nissenbaum, 1975; Oren, 2005; Spear *et al.,* 2006; Hallsworth *et al.*, 2007; Oren, 2010).

Furthermore, large diversity of unknown viruses may inhabit extreme environments. But unlike bacteria and archaea, there are limitations in finding universally conserved markers in viral genomes for further molecular studies (Edwards and Rohwer, 2005; FPaez-Espino et al., 2016). Moreover, viral quasi species are distributed across very different habitats and according to Kazlauskas *et al.,* 2019, caution was required when attributing viral genomes discovered by metagenomics to particular hosts. Hence, to minimize the concern raised due to limitations of culture-based and universal marker based techniques, metagenomic and metabolic analyses have been proven to be practical for studying microbiomes of extreme environments (Peña-Ocaña *et al.*, 2022).

2.2. Extreme Environments of Danakil Depression

2.2.1. Dallol Solfatra

The locally called “Dallol mountain” is an elliptically shaped mound rising nearly 50 m above the floor of the salt covered depression on top of the sub-aerial Dallol volcano. The sub-aerial Dallol volcano is a basaltic intrusion within the evaporites, slowly pushing up the overlying thick evaporite deposit, leading to the formation of a 1.5 km by 3 km “Dallol mountain” (Franzson *et al.*, 2015; Asrat, 2016; López-García *et al.*, 2020). Magmatic eruptions to the surface are seldom but numerous poweful phreatic eruptions have been recorded. For instance, the 1926 eruption which formed a 30 m wide crater at the top of the mound; the 2014 phreatic eruption of bischofite (hydrous magnesium chloride) rich hot brine, just southwest of the Dallol summit, as well as the minor phreatic eruption of January 2015 at the summit crater, are well recorded (Nobile *et al.,* 2010; Asrat, 2016). These phreatic eruptions, particularly the most recent eruption of 2015 led to the drying up of many brine/acid pools and the appearance of a few new ones at the summit crater (Asrat, 2016).

The circular, shallow depression at the summit of the Dallol mountain is believed to be a collapse crater, and nearly permanent, hyper-acidic and hyper-saline hot brine springs and pools on this crater are manifestations of an on-going volcanic-hydrothermal activity beneath the summit (Asrat, 2016; López-García *et al.*, 2020).

The acid pools are strongly variegated where the active, bubbling springs and pools are white to yellowish green and blue-hued deep green, while the inactive, dried up pools and springs are covered by reddish, brownish, yellowish and rust-coloured crusts (Franzson *et al*, 2015; Asrat, 2016; López-García *et al.*, 2020).

2.2.2. Black Water and Gaet'ale ponds

Besides Dallol Solfatra, other hydrothermal-geothermal activities exist in the southeast and southwest to the Dallol summit crater. Black Water pond, located about ~2 km southwest of mount Dallol in an area locally called "Black Mountain". The name 'Black' is attributed to the dark basaltic lava flow partly exposed beneath the salt deposits (Franzson *et al.*, 2015; Asrat, 2016). On the other hand, Gaet'ale is an actively bubbling, oily, sulphur-rich spring/pool spring, located ~4 km southeast of the Dallol Mountian. The Gaet'ale pond, similar to the Black water pond, originated from reactivation of a thermal spring during the 2005 major volcano-seismic activity in the Danakil Depression (Gebresilassie *et al.*, 2011; Franzson *et al.*, 2015; Master, 2016; Pérez and Chebude, 2017).

2.2.3. Lake As'ale and Mud'ara Pond

As'ale is a location inside the Danakil Depression, few kilometers south Dallol volcano, where of the hyper saline lake named Lake As'ale (also known as Lake

Karum or Assale) as well as number of lacustrine ponds reside (Abbate *et al.,* 2015; Hughes and Hughes, 1992). Lake As’ale is a very shallow, hyper-saline lake that lies 120 m below sea level (Hughes and Hughes, 1992; Abbate *et al.,* 2015; Asrat, 2016; Cavalazzi *et al.,* 2019).

Lake As’ale is believed to have been formed after the auxiliary sinking of the Danakil Depression (Beyth, 1978; Hughes and Hughes, 1992), which caused periodic flooding of the Red Sea that led to the accumulation of deep layers of marine salts several hundreds of meters thick. The Lake was part of the Red Sea until a volcanic barrier flow separated them less than 32,000 years ago (Beyth, 1978) and marine fossils from this area have been dated to 200 to 24,000 years before present (BP) (Lalou *et al.,* 1970; Abbate *et al.,* 2015). Runoff from the highlands to the west intermittently feed the Lake and the groundwater (Hughes and Hughes, 1992; Abbate *et al.,* 2015).

For centuries Lake As’ale has been exploited for its salt deposits estimated to be 2,200 m thick (Abbate *et al.,* 2015; Woldekiros, 2019). The salt deposit is composed of salt crusts varying from almost snow-white to dark brown, where the salt is mixed with dust and soil (Asrat, 2016). These salt layers represent the floor of a much larger fossil lake dated to 7-10 thousand years (Hughes and Hughes, 1992). Salt has been mined for centuries from the margins of Lake As’ale, a tradition that is still actively practiced by local workers (Woldekiros, 2019). The salt slabs are transported by camels, mules and donkeys (the ‘salt caravans’), to the western highlands through the harsh Danakil Depression and across the rugged escarpment rising from 120m below sea level to more than 2400 m above sea level (Asrat, 2016). The Lake is also dotted in many places by active salt diapirs with colorful, bubbling cold springs (Fig. 4). In stark contrast, Muda’ara Pond is a cold spring at the northern margin of the lake with no significant recorded natural history (Asrat, 2016).

2.3. Next Generation Sequencing (NGS)

For the first time, DNA was sequenced in a consistent and well defined way using Sanger chain termination method. Before Sanger method, sequencing was done using radioactive levels and two-dimensional chromatography methods that are capable of sequencing up to 15 nucleotides (Jay *et al.*, 1974). The automated version of the Sanger was coined the “First Generation Sequencing”, integrating fluorescently labeled terminators, capillary electrophoreses and automated laser signal detection. Using this standard, instruments producing ~84 kilo bases per run were introduced in the late 20th century by Applied Biosystems Company and were used for Human Genome Project (Liu *et al.*, 2012[b]). The NGS methods were developed in the early 21st century by considering the “massively parallel sequencing of short reads” approaches, where currently 1.8 tera bases of data can be generated in a single run using an Illumina HiSeq X Ten sequencer (Rizzo and Buck, 2012) (https://illumina.com/technology/next-generation-sequencing.html.)

Sequencing of the bacterial 16S ribosomal RNA gene was the first culture independent approach (Handelsman *et al.*, 1998). But comparing to the WMS approach, 16S rRNA gene sequencing is considered as low throughput since it only depends on one gene type for phylogenetic survey, while the former can be used to access the functional gene composition of microbial composition (Thomas *et al.*, 2012; Segata *et al.*, 2013). In addition, though 16S rRNA is one of the best marker genes for predicting genome-wide level gene sequence similarity for distantly related prokaryotes, it is not suitable for closely related ones (Lan *et al.*, 2016).

There are different types of platforms that use NGS technologies such as MinION, PacBio and NovaSeq (Baker, 2012; Rhoads and Au, 2015; Davies, 2017; Tyler *et al.*, 2018). But the 454-Roche and the Illumina/Solexa (HiSeq, NextSeq, and MiSeq) are the two systems of the NGS technology that have been extensively used for metagenomics studies, especially in microbial community studies (Luo *et al.*, 2012; Thomas *et al.*, 2012; Segata *et al.*, 2013; Luk *et al.*, 2015). These two platforms are reliable for quantitatively assessing diversity among natural communities and they similarities in the resulted estimated gene and genome abundance as well as in more than 90% of the assembled contigs and 89% of non-assembled reads (Luo *et al.*, 2012; Segata *et al.*, 2013). While Sanger sequencing is still considered the gold standard for sequencing, for its low error rate, long read length (> 700 bp) and large insert sizes which improve the outcome of the shot gun sequence, the NGS technology is the one that has dramatically accelerated the development of sequence-based metagenomics because of rapid and substantial cost reduction (Oulas *et al.*, 2015; Thomas *et al.*, 2012).

2.4. Development of metagenomic NGS (mNGS)

In microbiology, the conventional identification and characterization using morphology, staining properties and metabolic criteria have been more or less superseded by the genomic explanation of the particular microbe (Behjati and Tarpey, 2013; Cowan *et al.*, 2015). Since the introduction of NGS in the early 2000, genomic research activities have been revolutionized in a way that an entire organism's genome can be sequenced within days at minimum cost than Sanger sequencing (Behjati and Tarpey, 2013; Barba *et al.*, 2014; Selvarajan *et al.*, 2014; Cowan *et al.*, 2015). Unlike the Sanger sequencing technology which applies known primer sequences to extend into unknown nucleic acid strand, prior knowledge of the target sequences is not an absolute requirement in NGS, hence is performed de novo (Behjati and Tarpey, 2013; Barba *et al.*, 2014; Selvarajan *et al.*, 2014; Cowan *et al.*, 2015).

Sequencing of entire nucleic acids in a sample that contain mixed populations of microorganisms using any one of the several high-throughput sequencing methods (HiSeq, NextSeq, MiSeq, etc…) is known as metagenomic Next Generation Sequence [mNGS] (Bragg and Tyson, 2014; Gu *et al.,* 2019). Performing mNGS does not require any cultivation or enrichment for pure culture selection thus, technical biases raised due to culturing and isolating microbes are greatly reduced (Behjati and Tarpey, 2013; Cowan *et al.*, 2015). The ability to simultaneously sequence and understand, which microbes in what proportions are present within a sample makes mNGS a compelling platform for determining genetic materials from entirely different groups of organisms (Bragg and Tyson, 2014; Gu *et al.*, 2019).

Sometimes, mNGS incorporates polymerase chain reaction (PCR) (either targeted, universal or broad-range), which relies on specific primers such as 16S rRNA gene or internal transcribed spacer (ITS) for amplifying and detecting specific targets. On the contrary, shotgun mNGS is an impartial hypothesis-free method, where genome of mixed populations of microorganisms split into multiple fragments that are at the same time and independently sequenced (Bragg and Tyson, 2014; Hbarbalz *et al.,* 2016; Gu *et al.,* 2019). To perform this type of de-novo sequencing, paired end DNA libraries are often preferred as gap distance of each DNA fragment of the two yielded reads is known. The added information gained by knowing the distance between the two reads can be invaluable for spanning repetitive regions and enable adjacent regions of DNA to be ordered.

By and large, any approach of sequencing is only as good as the experimental design and DNA extraction protocols. Major drawbacks of mNGS are mostly related to identifying contamination, selection and validation of databases used for analyses (Bragg and Tyson, 2014; Gu *et al.*, 2019).

2.5. Data mining in metagenomics

According to Raza, (2010), data mining is defined as "the process of discovering meaningful new correlations, patterns, and trends by digging into large amount of data stored in warehouses". The main purpose of mining biological data is to extract useful knowledge from the bulky data set gathered in life sciences and make it into a reasonable form, especially in modern biology, where bioinformatics is applied to learn from the genetic data (Raza, 2010; Chitra and Maheswari, 2017). Supervised and unsupervised are two categories of learning from data and in the case of the latter, no variable is selected as a target, thus establishing some type of relationship among the variables is the goal (Raza, 2010; Chitra and Maheswari, 2017). Classification, estimation and prediction of data are examples of supervised learning while rules of association, clustering, description and visualization are unsupervised learning.

Investigation of microorganisms using metagenomics can be grouped into sequence based taxonomy profiling and function-driven (Selvarajan *et al.,* 2014; Cowan *et al*., 2015). Different bioinformatics tools can be applied for data mining, analysis and interpretation of the metagenomics data (Raza, 2010; Yue *et al*., 2020). The quality of result of sequence analyses such as marker gene analysis, sequence read assembly, and binning (i.e. downstream processes) is determined by the proper upstream processes of community DNA extraction, DNA fragmentation and sequencing.

2.6. Post sequence management of short reads

Sequencing is the primary step for obtaining raw data of genetic materials from environmental samples such as animal gut, soil, ocean, wastewater, hot springs, etc..., followed by genetic assembly (taxonomic and functional annotations) (Hill *et al*., 2000; Allaband *et al*., 2019; Cheng *et al*., 2019; Liu *et al*., 2020[a]). Current DNA sequencing technologies cannot read whole genomes in one run. As result, DNA got fragmented and sequenced, there are two approaches that should be considered (Ibarbalz *et al*., 2016; Segata *et al*., 2013). Grouping reads or contiguous sequences (contigs) in metagenome and assigning them to individual genome is known as metagenomic binning. Binning can be performed either on contigs or raw reads. The two types of binning approaches are taxonomic-dependent (taxonomy binning), and taxonomic-independent (genome binning). The first one involves reconstruction of the genome by arranging and merging the short reads of the fragmented DNA in a line and this step is called genome assembly (Ghurye *et al*., 2016; Howe *et al*., 2014; Luk *et al*., 2015). The second approach does not involve assembling of the fragmented DNA, rather uses the quality checked short segmented DNA reads as they are for binning (Segata *et al*., 2012; Segata *et al*., 2013; Sharpton, 2014; Pasolli *et al*., 2016).

To exclusively answer the question “who is there?” in microbial communities without considering the functional annotation, amplicon sequencing using specific marker genes or taxonomy profiling from short reads of mNGS can be sufficient (Sczyrba *et al*., 2017; Cheng *et al*., 2019; Yue *et al*., 2020). Based on the relative cost effectiveness, amplicon sequencing is an ideal strategy to have a preliminary understanding of the composition of the targeted microbial community (Cheng *et al*., 2019; Yue *et al*., 2020). On the other hand, the shotgun mNGS approach can address not only the question of “who is there in the community?” but also “how are they function as part of the community?” (Sczyrba *et al*., 2017; Cheng *et al*., 2019). Some values considered for utilizing shotgun mNGS are; (1) the high capacity of profiling taxonomic composition of microbial communities including low abundance and/or unknown microorganisms, (2) the potential of recovering partial or whole draft genome of microbes in the sample, and (3) the possibility of predicting how all microorganisms are networking in the communities through meaningful functional annotations of sequences (Sczyrba *et al*., 2017; Cheng *et al*., 2019; Yue *et al*., 2020).

The major limitation of mNGS sequencing, either targeted or random (shotgun mNGS), is the complexity of correctly and fully assembling short reads without gaping or completely missing fraction of reads (Alkan *et al*., 2011). Especially, when multiple strains are present in the metagenome sample, the complexity of sequence assembly increase exponentially (Sczyrba *et al*., 2017; Cheng *et al*., 2019).

2.6.1. Metagenome assembly and binning

Theoretically, reconstruction of complete genomes of organisms in a community can be achieved if the metagenomics sequence is done at sufficiently high coverage. However, in practice, in most cases whole-genome assembly from metagenomes is impossible, because of the intricate natural community composition where horizontal gene transfer, organismal variability and conserved DNA regions lead to non-resolvable uncertainties in scaffold reconstruction (Sczyrba *et al*., 2017; Segata *et al*., 2013).

Errors in assembly often occur because either pieces are incorrectly discarded as mistakes or repeats or they are joined up in the wrong places or orientations (Baker, 2012; Ghurye *et al*., 2016; Sczyrba *et al*., 2017). Lately, it has become somehow easy to study abundant organisms in novel complex communities or low to medium complex microbiomes, using assemblers specific for metagenome sequences without relying on previously sequenced isolates (Meyer *et al*., 2008; Narasingarao *et al*., 2012; Pell *et al*., 2012; Segata *et al*., 2013; Mikheenko *et al*., 2016). This type of assembling is termed as De novo assembling which effectively reduces the total amount of data to be analyzed but requires substantial computational resources (Howe *et al*., 2014). Furthermore, the new De novo metagenome assemblages are often evaluated on the basis of; the number of scaffolds and contigs required to represent the genome, the proportion of reads that can be assembled, the absolute length of contigs and scaffolds, and the length of contigs and scaffolds relative to the size of the genome (Baker, 2012; Narasingarao *et al*., 2012).

Three strategies namely, Greedy, Overlap-Layout-Consensus (OLC), and De Bruijn graph can be considered when handling de novo assembling with their strength and weakness (Ghurye *et al*., 2016). Greedy assembling strategy is the most simple and spontaneous method where individual reads are joined together into contigs starting with the reads that overlap best and ending once no more reads or contigs can be merged (Ghurye *et al*., 2016). The de Bruijn graph assembly method rely on relationship between k-mers (substrings of fixed length k) derived from the reads which are organized in a graph structure and reads are not explicitly aligned to each other, rather their overlaps can be concluded from the fact that they share k-mers (Pell *et al*., 2012). However errors in the reads can introduce false k-mers (false nodes and edges) in the graph hence need to be eliminated before identifying an Eulerian path in the graph (Ghurye *et al*., 2016; Pell *et al*., 2012). Overlap-Layout-Consensus approach is also a graph based method which involves finding overlaps between all parts of read before constructing the graph. Contigs are constructed by finding branchless paths in the graph and get further organized and extended using mate pair information. The overlapping reads and the corresponding paths in the graph must show consensus on the sequence before the contigs are considered for assembly (Ghurye *et al*., 2016).

Progress in metagenomic assemblers based on de Bruijin graph make improvement of the length of contigs or scaffolds and the number of predicated genes and incorporated sequences (Compeau *et al*., 2011; Tran and Phan, 2020). Not only can long contigs or scaffolds with less error by utilizing modern assembly tools can reduce the loss of sequence features but also make employing the co-abundance of taxon across multiple samples possible in genome binning.

Bioinformatic tools such as Phylopythia (McHardy *et al*., 2007) and MEGAN6 (Huson *et al*., 2016) are examples of taxonomy binning approach. However, MEGAN6 is alignment-based binning algorithm that use Double Index Alignment Of NGS Data (DIAMOND) (Buchfink *et al*., 2015) against protein reference database, such as National Center for Biotechnology Information non redundant (NCBI-nr) protein database. Binning assembled sequences into individual groups enables the functional annotations of the sequences to be related in a meaningful way (Alneberg *et al*., 2014; Yue *et al*., 2020). Like in any data mining processes, both supervised and unsupervised machine learning methods can be employed in metagenome binning (Alneberg *et al*., 2014; Yue *et al*., 2020). Bioinformatic tools such as Phylopythia (McHardy *et al*., 2007) are used for supervised metagenome binning.

2.6.2. Assembly-Free Taxonomic profiling methods

As the length of sequence reads get shorter and assembling becomes close to impossible, several alternative methods of taxonomy profiling have been developed. Such methods are designed to rely on the extraction of only the most informative features from reference genomes to reduce the complication of mapping processes (Segata *et al*., 2012; Segata *et al*., 2013; Koslicki *et al*., 2014; Truong *et al*., 2015). Many of the profiling tools that perform without assembling sequence reads are extrinsic. These bioinformatic tools rely in some way on reference genome catalogs for comparisons of metagenomic sequences with reference sequences in order to identify taxonomic or phylogenetic origin (Segata *et al*., 2013; Darling *et al*., 2014; Koslicki *et al*., 2014; Chen *et al*., 2017). However, reference-independent approaches that use key sequence properties for composition-based clustering can also be adopted for intrinsic (de novo) binning. Bioinformatics tools that apply reference- independent approaches do significantly better for environments with insufficient prior information of genomic but consume extended computational times (Patil *et al.*, 2012; Segata *et al.*, 2013). Nonetheless, both intrinsic binning and/or extrinsic (homology reference-based) approaches help classification of metagenome shotgun reads and provide comparable information to investigate complex communities.

Profiling microbial communities using universal markers such as 16S rRNA genes may improve generalizability and phylogenetic resolution of the metagenome analysis and can help estimate any sufficiently sized sequence sample (Patil *et al.*, 2012; Segata *et al.*, 2012; Gregor *et al.*, 2016; Lan *et al.*, 2016). However, integrating shotgun metagenome sequencing with analysis of pre-identified unique clade-specific marker genes (considered as tags) to help discriminate different strains within species and allow more precise organismal abundance estimation (Patil *et al*., 2012; Segata *et al.*, 2012; Darling *et al.*, 2014; Truong *et al.*, 2015; Gregor *et al.*, 2016; Hbarbalz *et al.*, 2016; Lan *et al.*, 2016). Therefore, focusing on unique clade-specific marker genes can help to cut back the computational run times (Segata *et al.*, 2012; Truong *et al.*, 2015). However, most homology reference-based methods need to incorporate complete genome searches against the entire catalog of reference genomes or mapping for binning and community profiling. The advantage of whole genome searching is the availability of information on sequence variation of individual microbes as well as gene and pathway collections of a community. But this type of approach can be highly difficult to interpret and cause uncertainty with the false positive results generated because of evolutionary conserved or horizontally transferred sequences. Yet different types of computational strategies (such as alignment of reads to the lowest common ancestor of taxa with hits) have been designed to correct the ambiguity caused by complete genome lookup and provide accurate taxonomic characterization.

2.7. In silico analysis of metagenome data to investigate microorganisms and bio catalysts from extreme environments

Applications of biocatalysts (whole cell or enzyme) in industries such as food and feed, pharmaceutical, textile, detergent, beverage and so on are becoming common practice (Littlechild, 2015; Cáceres-Moreno *et al.*, 2019). The global need for “green industry” for high specificity, selectivity and activity but produce zero toxic waste is the driving force for the growth of through “white biotechnology” (Littlechild, 2015; Cáceres-Moreno *et al.*, 2019; Salwan and Sharma, 2019). However, some enzymes can also undergo unwanted side reactions when used at a different physical conditions inside bioreactors (eg. pH) compared to the inside of the source cell (Littlechild, 2015). Hence, investigation of extremozymes from different environments provide the opportunity to access stable enzymes in extreme temperatures, salt concentrations, pressure, pH, or combination of these properties (Littlechild, 2015; Cabrera and Blamey, 2018; Cáceres-Moreno *et al*., 2019; Salwan and Sharma, 2019).

Extremozymes such as transaminases, carbonic anhydrases, dehalogenases, specific esterases, and epoxide hydrolases have been studied using bioinformatic techniques and screening of expression libraries to carry out a variety of chemical biotransformations of commercial interest in different industries (Littlechild, 2015).

The natural trove for extremozymes are microorganisms living in **e**xtreme environments are believed to be representatives of the oldest inhabitants of Earth (Selvarajan *et al.*, 2014; Cowan *et al.*, 2015). Understanding the evolution, biogeochemistry, pathogenesis, and genetic determinants underlying the biotechnological potential of extremophiles is very challenging as majority of microorganisms on Earth cannot be cultivated and characterized using conventional laboratory methods (Selvarajan *et al.*, 2014; Cowan *et al.*, 2015). Regardless of the culturability of microbes, genomic data obtained directly from an environmental sample can provide knowledge on how inhabitant species function and interact with in communities of the microhabitat.

The expression “in silico” in biology means characterizing all biological experiments carried out in a computer (Miramontes, 1992; Barnard, 2014). The progress in different bioinformatics tools has made the “in silico” prediction and design of new functional molecules possible (Barnard, 2014; Dana *et al.*, 2020). In environmental microbiology research, there is a general analytical progression with increasing technical and system complexity that first addresses community diversity and then environmental drivers of the community followed by community function study (Cowan *et al*., 2015). Metagenomics can be applied exclusively or coupled with other "omic" technologies (transcriptomics, proteomics and metabolomics) to improve the collective analysis of microbial community structure as well as the inter and intra community functional interactions within and/or with the environment (Selvarajan *et al*., 2014; Cowan *et al*., 2015; Kleiner *et al*., 2017; Fahrenfeld *et al*., 2017; Perito and Cavalieri, 2018; Feng *et al*., 2018; Bedoya *et al*., 2019).

Currently, there is lack of consensus about bench-marking performance assessment of methods applied in shotgun metagenomics (Sczyrba *et al*., 2017; Yue *et al*., 2020). Some may use taxonomy profiling software like MetaPhlAn 2 (Segata *et al*., 2012) to predict the taxonomic identities and relative abundances of microbial community members (Fig. 2). On the other hand, others may prefer taxonomic binning software such as MEGAN6 (Hudson *et al*., 2016) to classify contigs or reads with taxonomic label attached and ideally generating draft genomes of a strain (or higher-ranking taxon) from a microbial community. Therefore, it is critical to have guideline for software selection to answer specific research questions in metagenomics (Sczyrba *et al*., 2017). However, generally applied key methods for In silico interpretation and understanding of metagenomic data are assembly, taxonomic profiling and binning (Sczyrba *et al*., 2017).

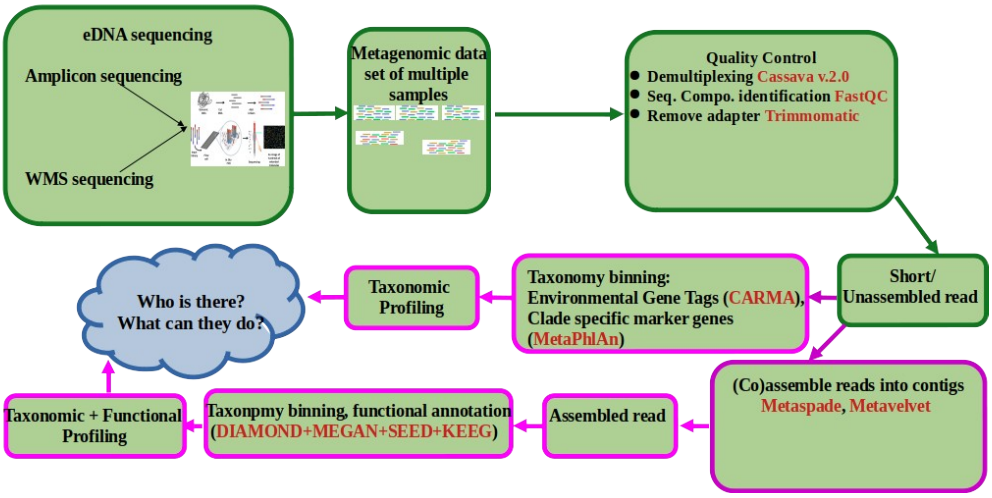


Fig. 2: Options of metagenom sequence analysis

There are four in silico approaches that can be used to predict protein coding sequences and metabolic pathways from mNGS reads of microbial community. First, custom database of protein-coding or clade specific genes can be constructed from the set of reference organisms detected in the metagenome sample (e.g. HUMAnN2, MetaPhlAn2) (Segata *et al.*, 2012). Second, de novo assembly is done to combined short reads from mNGS into contigs, which can be further used to identify open reading frames (e.g. metaSPAdes, IDBA-UD) (Peng *et al.*, 2012; Nurk *et al.*, 2017). Third, assembled reads can be directly mapped (aligned) to a closed reference of protein coding sequences (MEGAN 6, DIAMOND). Forth, based on the family of proteins predicted, metabolic pathways can be mapped (KEGG, SEED) (Kanehisa and Goto, 2000; Bassel *et al.*, 2011).

Chapter Three

3. Diversity of Prokaryotes In Hypersaline Lake And Acidic Ponds Of Danakil Depression Using Next Generation Whole Metagenome Shotgun Sequence Analysis

Abstract

The Danakil Depression in northern Ethiopia is a place with numerous polyextreme environments located. These poly-extreme environments exceed the normal physico-chemical limits of either pH, salinity, ion content, temperature or all. In this study, brine samples were collected from five different locations in the Danakil Depression; the coastal parts of Lake As’ale, an active salt diapir pond known as Mud’ara (MUP) and the acid ponds of Dallol (DAL), Black Water (BLA) and Gaet’ale (GAL). The presence of DNA-based life in these environments were tested using Whole Metagenomic Shotgun (WMS) DNA sequencing approaches. The phylum Euryarchaeota accounted more than 95% of the total Operational Taxonomic Unit (OTU) in Lake As’ale and MUP. For DAL, GAL and BLA, more than 90% of the identified OTUs belonged to Bacteria, mainly to phylum Proteobacteria. From the total profiled OTUs in DAL, GAL and BLA, more than 79% were unclassified at different taxonomic levels. For coastal parts of Lake As'ale and MUP, little more than 50% of the OTUs were classified at species level and *Halonotius* sp *J07HN6* and unclassified *Salinibacter ruber* were highly abundant and shared archchaea and bacteria OTUs, respectively. Unclassified strain of *Paraburkholderia fungorum* was the most abundant profiled OTU for DAL and BLA, while unclassified *Escherichia* and unclassified *Burkholderia* were the first and second highly abundant OTUs respectively for GAL. Qualitative analyses of species richness and species diversity based on MetaPhlAn 2 results showed unique microbial composition in HsAP and HsAqS. The minimum spanning tree based on Bray-Curtis disimilarity distances indicated two clusters, the three sampling sites of the HsAP and the two sampling sites of HsAqS, respectively.

Keywords; Danakil Depression, Halophiles, Acidophiles, Polyextremophiles

3.1. Introduction

The extremity of environments is determined by considering the physico-chemical conditions such as acidity/alkalinity, salinity, temperature, high concentrations of heavy metals, and others. The East African Rift System (EARS), particularly the Afar Depression has been a focal point for investigating continental fragmentation and the first stages of oceanization (Abbate *et al.*, 2015). The Danakil Depression is located in the northern part of the larger Afar Depression (Fig. 3 A and B), and it is a segment of proto-oceanic crust where lowland plains occur 116-125 m below sea level and are split by fault blocks and dotted with shield volcanoes (Asrat, 2016; Cavalazzi *et al.*, 2019). This depression is one of the most arid places on Earth, with the highest annual average temperature reaching 50–60°C (120–140°F), very little rainfall, and intense sunlight (Beyth, 1978; Ahrens, 2011; Abbate *et al.*, 2015; Fazzini *et al.*, 2015; Vereb *et al.*, 2020).

Danakil is dotted with volcanos including the active Ert’ale and the quasi-active, sub-aerial Dallol, buried under a relatively thick sediment cover (Asrat, 2016). The central North West-South East rift axis of the depression joining Dallol and Ert’ale marks an active sea floor spreading zone leading to an eventual continental breakup, as is evident from the intense earthquake and volcanic activities focused along the same axis

East (NW-SE) rift axis of the Depression joining Dallol and Ert’ale marks an active sea floor spreading zone leading to an eventual continental breakup, as is evident from the intense earthquake and volcanic activities focused along the same axis (Illsley-Kemp *et al.*, 2017).

Fig. 3 Satellite image of location of Danakil Depression

Location of Danakil Depression on Map of Ethiopia (A) and in Afar Triangle (B). *(Source: https://maps-ethiopia.com/danakil-ethiopia-map).*

Active volcanism beneath the evaporites makes Danakil a uniquely colourful place. Succession of thick evaporitic deposits (halite and other salt deposits intercalated with frequent but thin carbonate and clastic sequences), which currently cover the entire floor of the depression, were formed due to the subsequent drying up of periodically transgressing marine water from the Red Sea during the late quaternary (Asrat, 2016). The high deposit of salts that makes up the entire floor of the valley and the active volcano in addition to the intense climatic conditions contributed a lot for the otherworldly features of the depression (Asrat, 2016). In general, Danakil, has an extraterrestrial appearance and quality, which could be rightly considered inhospitable to living things. Particularly, the acid pools found in the depression are ideal places to study how the poly-extreme environments affect organic life on Earth or other sites in the solar system, since life is believed to have begun at the aqueous surface of sulfide minerals and acidic pH (Wächtershäuser, 2006; Barbieri and Cavalazzi, 2014, Vereb *et al.,* 2020).

Among the several reasons that can warrant studying diversity of microorganisms in extreme environments, the polyextreme conditions of the Danakil Depression were major drives for studying diversity and biotechnological potentials of polyextremophiles (Cavalazzi *et al.,* 2019). Unlike the Soda lakes and hot springs found in the main Rift valley system of Ethiopia (Kebede *et al.*, 1994; Lanzén *et al.,* 2013; Simachew *et al.*, 2016), little is known about the biological resources of the extreme habitats in Danakil. The integration of our results with those from extensively studied soda lakes of the main Rift valley system of Ethiopia can give valuable information on how early or other planetary microbial life forms adapted and evolved in early evaporitic environments with high concentrations of salt (Jiang *et al.*, 2006 ; Barbieri and Cavalazzi, 2014; Li *et al.*, 2014).

Thus, in this report, the authors present results from comprehensive field surveys conducted on the sample sites in Danakil Depression in 2015 and the metagenome analysis of environmental samples taken from the sites. The goal was to determine the chemical characteristics of the samples from the study sites and the microbial diversity using molecular analysis. The results of the molecular analysis, which were based on WMS sequencing using Illumina next generation platform, were further substantiated by non assembled bioinformatic tools to evaluate variation and diversity of microorganisms in the study sites.

3.2. Materials and Methods

3.2.1. Sample collection

The Danakil Depression is generally located on the Northern Hemisphere, at 37 longitudinal zone. As Lake As’ale is a large saline lake on a white salt plain at -120 m below sea level (Fig. 4), brine samples were collected from three different locations on the coastal parts of the lake and labeled as; LA (UTM 0646610E, 1558846N), LADP5 (UTM 0648365E, 1558709N) and LADP6 (UTM 0645195E, 1558770N) on 3rd and 4th of February, 2015. For Muda’ara pond (MUP), brine samples were collected at UTM location 0645190E/1558764N, -109 m below sea level on 4th of February, 2015 (Fig.4).

In the case of Dallol solfatra, due to the phreatic eruption which led to the draining and drying up of most of the active springs at the crater, only one pond (DAL) located at UTM 0639983 Easting/1574575 Northing was available for sampling on 4th of February, 2015 (Fig. 5). Collection of brine samples from Black Water (BLA) and Gaet’ale (GAL) was done on the same day at locations UTM 0638790 Easting/1572643 Northing and 0642556 Easting/1571715 Northing respectively (Fig. 5). All the brine samples were collected randomly and in triplicates from accessible parts considering minimal human and animal contacts to reduce chances of contamination. Photos of sample sites are shown on Fig. 6.

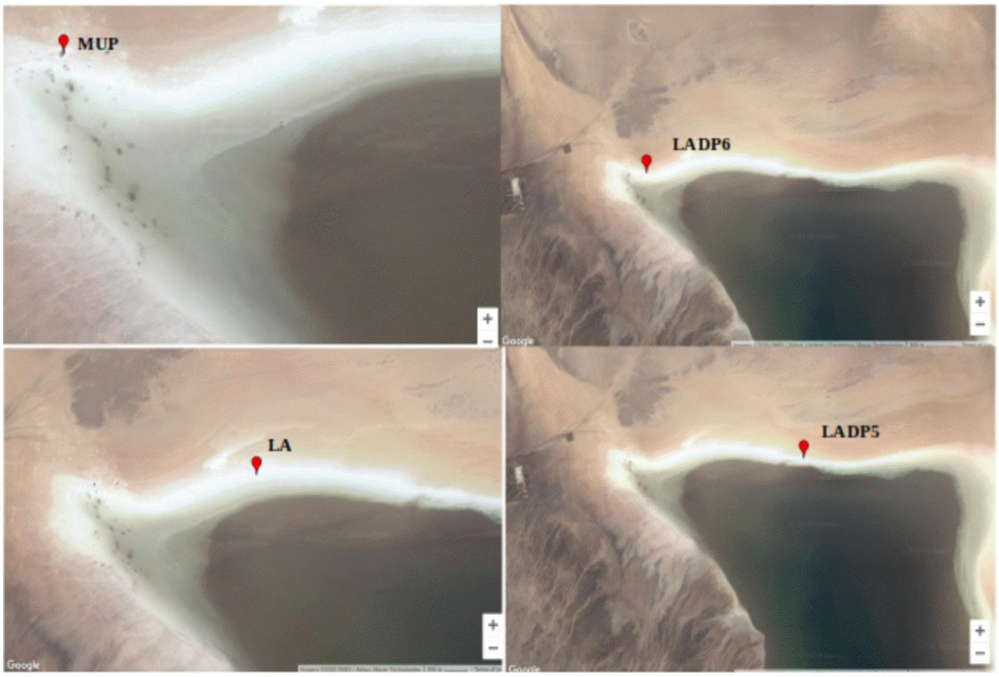


Fig. 4: Satellite images of sampling sites on Lake As’ale and Mud’ara pond.
(Source: Google Earth)

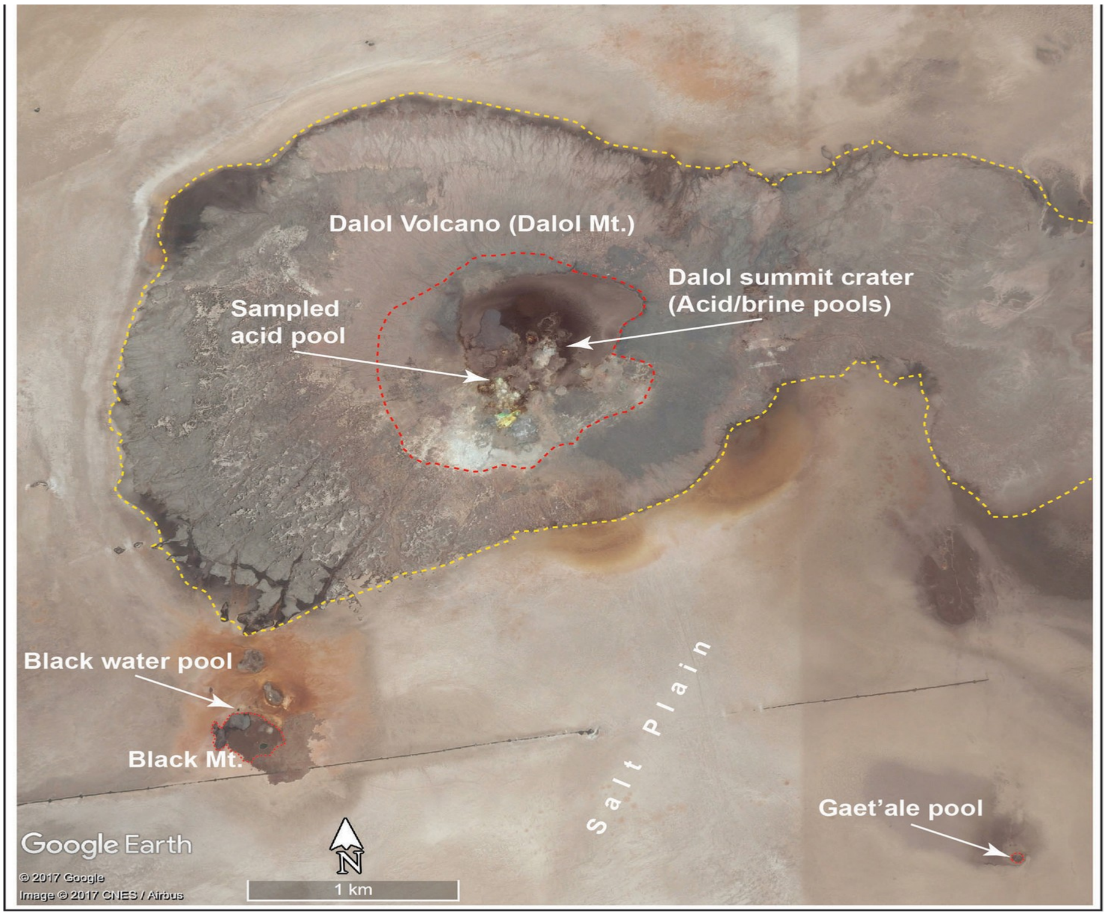


Fig. 5 Satellite image of Dallol, Gaet'ale and Black Water.

(*Source: Google Earth)*

Newly formed spring on top of the Dallol crater (A); Gaet'ale (B); Lake As'ale (C); Mud'ara (D) and Black water pond (E). (Photo by Lulit, 2015)

Measurements of pH, conductivity and temperature were done onsite for coastal parts of Lake As'ale and MUP using portable pH and conductivity meter (430 Enterprise Portable pH and Conductivity Meter 430271, Jenway). Salinity was measured using a refractometer (DIGIT-0120 ATC, VWR, UK) after diluting the brine samples 1 to 10 times with deionized water. However, onsite measurements of pH, temperature and conductivity could not be done for DAL, BLA and GAL due to the unpredictable and unsafe conditions for the measuring devices as well as for prolonged stay on the sample sites after the phreatic eruption at the Dallol summit crater. Hence, measurement of pH was done using 0-14 pH indicator paper sticks (Fisherbrand™) while insitu measurements of temperature and conductivity were avoided. Salinity was measured using a refractometer (DIGIT-0120 ATC, VWR, UK) by diluting the brine samples 1 to 10X with deionized water. The color, texture and other physicochemical characteristics of the samples are listed in Table 1.

3.2.2. Sample preparation for metagenome and physico-chemical analysis

Brine samples for metagenome analysis were collected using sterile High Density Polyethylene (HDPE) bottles for cell harvesting. Additional 50 ml of brine samples were collected in conical plastic tubes (BD Falcon TM, Greiner Bio-One, Germany), for other biological and physicochemical analysis. The samples were stored in an icebox containing cooling elements and transported immediately to a temporary working space at Yara Dallol BV, Dallol Camp for processing.

A total of 2000 ml of brine sample (200 ml per filter) from four sample sites (DAL, GAL, Lake Asale [LA, LADP5, LADP6] and MUP) were filtered without dilution using a 0.22 µm GE® polycarbonate filter membranes. However due to high viscosity, only 500 ml of brine (50 ml per filter from BLA) was filtered without dilution using a 0.22 µm GE® polycarbonate filter membranes. Then, the filter membranes were stored in a 2 ml Lasany® internal treaded Cryo-vials with sucrose lysis buffer (0.75 M sucrose, 20 mM EDTA, 200 mM NaCl, and 50 mM Tris-HCl, pH 9.0) for preserving the cells.

For physicochemical analysis, 200 ml of the filtrate was placed in sterile containers without acid treatment. All the processed samples (the filter membranes for metagenomic DNA extraction and the filtrate brine for physicochemical analysis) were transported to the Microbial Biotechnology laboratory at Addis Ababa University in iceboxes containing cooling elements and then stored at -20 ^{0}C and 4 ^{0}C, respectively until further analysis.

3.2.3. Physico-chemical analysis and water isotope measurement

Stable isotope compositions of oxygen and hydrogen from sampling sites were determined at the School of Earth Sciences, Isotope Hydrology Laboratory, Addis Ababa University, using the Los Gatos Research Off-Axis Integrated Cavity Output Spectroscopy (OA-ICOS), following standard procedures as recommended by Emanuelsson *et al.* (2015). However, only partial hydrochemical analysis was feasible to conduct at the Laboratory of Addis Ababa Environmental Protection Authority. The content of magnesium (Mg^{2+}) was determined using an atomic absorption spectrometer and the graphite method after 10,000X dilution. Total phosphorus (TP), nitrate (NO_3^-), nitrite (NO_2^-) and sulfate (SO_4^{2-}) were assayed according to the standard methods for the examination of water and wastewater (APHA, 1998). The chemical oxygen demand (COD) was measured using a Hach COD reactor 45600, USA. Chloride (Cl^-) testing was done using argentometric titration where potassium chromate was used as indicator.

3.2.4. DNA extraction

Environmental DNA (eDNA) was extracted at the Microbial Biotechnology Laboratory (Addis Ababa University) and at the PrIMO Laboratory (Brown University, Providence RI) using a modified Cetrimonium bromide (CTAB) method obtained from Zhou *et al.* (1996) and Mitchell and Takacs-Vesbach (2008). Considering the extremity of the environments where the sample taken, intensive optimization for extracting, cleaning, and concentrating of DNA was needed. For all samples of Lake As’ale and MUP pond, eDNA extraction was accomplished by directly following the optimized CTAB method from Zhou *et al.* (1996). For GAL sample, the 1% CTAB-Sodium Dodecyl Sulfate (SDS) DNA extraction method from Mitchell and Takacs-Vesbach, 2008 was used for optimum eDNA extraction. This method was proved to be ideal for extracting eDNA from acidic hot springs at

Yellowstone National Park, USA. Nonetheless for DAL sample, modification on DNA extraction was required due to the reasons described in Appendix 1.

Therefore, the optimized eDNA extraction methods employed for Dallol acid pond is as follow:

1. 1% CTAB buffer solution was prepared (1% CTAB, 0.75 M NaCl, 50m M Tris pH 8, 10m M EDTA)
2. The 0.22 µm GE® polycarbonate filter membranes with the non toxic Sucrose lysis buffer was transfered into 50 ml Sterile, certified RNase-/DNase-free and nonpyrogenic conical tubes (Corning™ Centrifuge Tubes, USA)
3. 1% CTAB buffer solution was added in to the 50 ml tube and wash the filter membrane by vortexing
4. 2ml cell containing buffer solution was distributed into a 15 ml Sterile, certified RNase-/DNase-free and nonpyrogenic conical tubes (Corning™ Centrifuge Tubes, USA) and place them in Bioruptor® Sonication System (a cycling parameter of 30 seconds on and 30 seconds off for 10 minutes)
5. Proteinase K (10mg/ml) final concentration was added and incubate for 2 hour at 60 ^{0}C
6. 20% SDS (final concentration 2%) was added and incubate for 1-2 hours at 60 ^{0}C and then centrifuge at 6000 X g for 10min

7. Top phase of the supernatant was transferred into a 2ml micro centrifuge tube then equal amount phenol:chloroform was added and centrifuged at 4000xg for10 min
8. Top phase of the supernatant was transferred into a new 2ml micro centrifuge tube equal amount chloroform was added and centrifuged at 4000xg for10 min
9. Top phase of the supernatant was transferred into a new micro centrifuge tube then 0.6 volume of Isopropanol was added and placed at 4 ^{0}C overnight then centrifuged 16, 000 X g for 10 min at 4 ^{0}C
10. Wash the precipitated pellet with 70% ethanol and re suspended with Tris-EDTA (TE) buffer
11. DNA Clean and Concentrator™-5 (Zymo Research) was used to purify and concentrate the DNA per the manufacturers instructions

Since the amount of brine sample filtered from BLA pond was small due to high viscosity, the optimized extraction method for DAL sample was directly adopted for extracting eDNA from BLA. All eDNA extractions were performed in triplets. To minimize DNA extraction bias, the three replicate of extracted DNA were later pooled prior to metagenome sequencing. The quantity and quality of eDNA from all brine samples were checked using PicoGreen Assay and Thermo Scientific NanoDrop 3300 Fluorospectrometer.

3.2.5. Library construction and WMS Sequencing

DNA fragmentation was performed using a CovarisTM S220 Ultra-Sonicator and fragments sizes were checked using Fragment Analyzer™ Automated CE System and DNF-486-33 High Sensitivity NGS Fragment Analysis Kit (35 bp - 6,000 bp) at the Genomics Core facility (Brown University, Providence RI). Four dual indexed libraries were prepared using TruSeq NANO DNA LT library prep kit (Illumina, San Diego, CA, USA) per the instruction in the manual. Libraries were barcoded and combined into a single group and sequenced on one lane of a flow cell using a 100 bp paired-end run on a HiSeq 2500 instrument (Illumina) at the Genomics Core facility (Brown University, Providence RI).

3.2.6. Metagenomic sequence processing

The reads from sequence runs were demultiplexed using Cassava v.2.0 and the detected barcodes in the report were well balanced. Sequence composition of the raw data was identified using FastQC. Trimmomatic was used to remove any adapter contamination (Bolger *et al.*, 2014). The software settings for Trimmomatic were ‘ILLUMINACLIP:adapters.fa:2:30:10’ to remove the TruSeq adapter sequences and ‘LEADING:3 TRAILING:3 SLIDINGWINDOW:4:10’ to improve mean sequence quality by trimming leading and trailing bases with $Q < 3$ and any sequences with a four-base sliding window mean below Q20. The sequences were then aligned to the human (GRCh37) and the phage PhiX genome (a standard for Illumina sequencing) to find and filter any non-microbial contamination.

3.2.7. Taxonomic assignment of non assembled reads and statistical analysis

Metagenomic Phylogenetic Analysis (MetaPhlAn 2) bioinformatics tool was used for mapping direct sequence reads against a reduced set of clade specific marker sequences (Segata *et al.,* 2012; Truong *et al.*, 2015). The direct non assembled, quality-checked and merged sequence reads were input data for MetaPhlAn 2. Microbial relative abundance profiles were generated with MetaPhlAn2 (Segata *et al.,* 2012) using default parameters and bowtie2 alignment. The MetaPhlAn2 reference database consists clade-specific marker genes from ~17 000 reference genomes (79% bacteria/archeaa, 20.4% viral and 0.6% eukaryotic). The marker genes in the database are genes that are unambiguously characterized in a taxonomic clade as they are always present in the sequenced isolates of that clade and never present in any other sequenced organism.

Operational taxonomic unit (OTU) profiles from all studied samples were merged with the script ‘merge_metaphlan_tables.py’ included with the MetaPhlAn2. Distribution and heatmaps were generated with ‘metaphlan_hclust_heatmap.py’ script using default options and the ‘-d braycurtis’, ‘-minv 0.01’ flags. GraPhlAn tool was used to visualize microbial abundances on a tree of life by generating circular taxonomic and phylogenetic trees (Asnicar *et al.*, 2015). To generate the phylogenetic trees, GraPhlAn requires two input files: a tree file and an annotation file. The export2graphlan module coverts results of the analysis of MetaPhlAn2 to tree (‘tree.txt’) files and annotation (‘annot.txt’) files using the script ‘export2graphlan.py’. The two files are subsequently used with GraPhlAn’s script ‘graphlan_annotate.py’ to generate PhyloXML files that can be converted into images by the script ‘graphlan.py’.

Multidimensional scaling (MDS), also known as Principal coordinate analsysis (PCoA) of the sample abundance profiles was carried out by first merging MetaPhlan2 results into a single BIOM format table. A qiime utility script for calculating the Bray-Curtis (Bray and Curtis, 1957) dissimilarity matrix on BIOM-formatted taxonomic profiles was then used. The default emperor plot was used to build the PCoA plot.

The R Studio with Vegan and Mass packages were used for calculating similarity indices (Bray-Curtis index and Jaccard’s index), diversity indices (Shannon’s index annd Simpson’s index), non parametric estimation of number of species in a community (Chao1). The functions “vegdist” and “diversity” in Vegan package were applied to calculate the dissimilarity indices (Bray-Curtis index and Jaccard’s index) and diversity indices (Shannon’s index annd Simpson’s index) respectively. For calculating Bray-Curtis dissimilarity, the default for the function “vegdist” was changed to “binary =TRUE” (i.e. (b+c)/ (2a+b+c), where a denoted the number of species shared between two sites while b and c indicated unique species). Hierarchical clustering was plotted based on Bray-Curtis dissimilarity distance using function “spantree”, which finds a minimum spanning tree for dissimilarities connecting all points. R is a freely available program for statistical computation and graphics. It is available through local CRAN mirrors (http://cran.rproject.org/).

3.3. Results

3.3.1. Geochemical and physicochemical properties

The result of salinity analysis showed that high percentage of salinity in all sampling sites (Table 1). The percent salinity of GAL and BLA ponds were the highest (68% and 60% respectively) while for DAL, salinity was 34%. Insitu measurements of pH, temperature and conductivity indicated that the studied coastal sites of Lake As’ale and MUP were hyper saline (salinity 32% and 36% respectively), slightly acidic (pH 4.25-5.61) with warm water temperatures (≥30°C) (Table 1).

Table 1 Geochemical Data of the sample site

Sample sites labeling code	**GPS location**		**Altitude (meter)**	**Description of sampling sites**	**Ave. pH**	**Ave. salinity(%)**	**Ave. T^o (^{0}C)**	**Ave. EC (mS/cm)**
	UTM.EW	UTM.NS						

DAL	639983	1574575	-95	Light green/ oily	<1	34%	NA	NA
GAL	642556	1571715	-115	Deep orange/oily	<1	68%	NA	NA
BLA	638790	1572643	-114	Dusky/oily	<1	60%	NA	NA
Lake As’ale								
							32^{0}C	237
(LA)	0646610	1558846	-120	Large shallow lake on white salt plain	5.60	32%	32^{0}C	236
(LADP5)	0648365	1558709	-115	Large shallow lake on white salt plain	5.61	32%	31^{0}C	237
(LADP6)	0645195	1558770	-115	Large shallow lake on white salt plain	5.61	32.1%		
MUP	645190	15 58764	-111	Small mixing greenish color pond adjacent to LADP6	4.25	35.6%	30^{0}C	235

NA= Data not available

The outcome of measurement of ion contents showed distinct variation among the sample sites (Table 2). The Total Phosphorous (TP) recorded in DAL was the highest (1.2 g/l) while zero amount was measured in MUP and LA. The content of Mg^{2+} ion in the BLA (300.4 g/l) was the highest among all sample sites and the amount of Cl^- ions in all brine samples was comparatively high (> 200 g/l). The results of other measured ions (SO_4^{2-}, COD, NO_2^-, and NO_3^-) are listed in Table 2. The oxygen and hydrogen isotope analysis (Table 2) showed that the BLA and GAL are characterized by positive δ^2H but depleted δ^{18}O, attesting to the origin of the brine to deeply circulating groundwater generated from meteoric water from cooler and highland sources. On the other hand, the DAL is characterized by strongly enriched δ^{18}O and slightly depleted δ^2H. The brine samples from LA and MUP were characterized by enriched δ^{18}O and δ^2H and highly depleted values respectively.

Table 2 Selected hydro-chemical and stable water isotope measurements

Sites	**Mg^{2+}(g/l)**	**NO_2^-(mg/l)**	**NO_3^-(g/l)**	**TP(g/l)**	**COD**	**SO_4^{2-}(g/l)**	**Cl^-(g/l)**	**$\delta^{18}O$ (/mil)**	**δ^2H (/mil)**
DAL	3.70	1.94	0.76	1.20	12150	5.50	234.30	8.56	-5.89
GAL	19.50	0.12	0.11	0.45	17500	0.00	432.80	-5.99	15.83
BLA	300.40	0.00	0.20	0.13	14150	0.00	311.10	-6.22	45.17
LA	59.40	1.34	0.20	0.00	2450	0.40	217.30	11.51	46.36
MUP	7.80	2.40	0.130	0.00	1925	0.10	219.10	-1.38	3.42

3.3.2. Metagenome sequence analyses

All filtered brine samples yielded genomic DNA quantities exceeding the minimum required 1 ng for Illumina 2500 library preparation (Table 3) with an average library insert of 500 bp. The total number of reads generated from sequencing is reported in Table 3. The highest number of sequence reads was generated from GAL (29,741,784). DAL (3,480,089) and BLA (6,147,722) were less than sought due to intense optimization procedures of DNA extraction and cleaning. From the studied coastal sites on Lake As’ale (LA, LADP5, LADP6) and MUP total of 37,565,144 and 11,016,361 reads with average GC content of 55% and 61% were generated respectively. The average GC content of reads from DAL, GAL and BLA were ranged between 52-58%. However, the highest GC content was recorded for reads from MUP.

In general, the read quality for all samples was average in the leading and trailing ends of the reads, and very good at all positions within the reads and overall the sequence depth was ample to analyze in detail. After removing the adapter sequences and checking the quality, all reads were 100 base pair (bp) long (Table 3). Direct shotgun mNGS data was used by MetaPhlAn2 to generate merged taxonomy profile table of microbial communities in the studied sample sites (Appendix 2).

Table 3 Overview of Metagenomics

Sites	**Sequence ID**	**Estimated DNA conc. for sequencing (ng/µl)**	**Length of single read (bp)**	**Total number of reads**	**GC content (%)**
DAL	LTW0001	> 0.8	100	3480089	54
GAL	LTW0002	2.128	100	29741784	52
BLA	LTW0003	> 0.7	100	6147722	58
LADP5	LTW0004	79	100	13103002	51
LADP6	LTW0005	52	100	14021795	56
LA	LTW0006	137	100	10440347	53
MUP	LTW0007	58	100	11016361	61

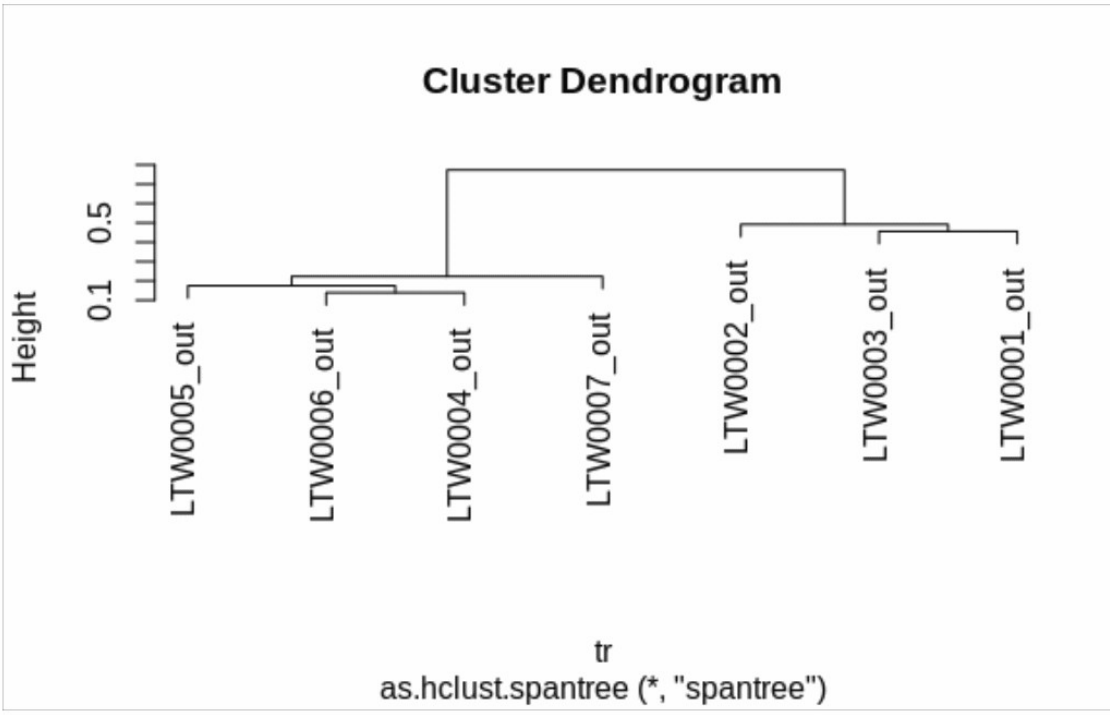
3.3.3. Statistical analysis and taxonomic profiling

All calculated indices of species richness and species diversity based on MetaPhlAn 2 result are shown in Table 4. Species diversity in a given community is measured with diversity indices such as Shannon or Simpson. Remarkably, sites in DAL, GAL and BLA harbored different microbial community compositions, even if the number of observed OTUs is moderate (average of 50 individuals). The Shannon index showed the highest values in these samples (range between 4.0 and 4.2) and the lowest ones in all other HsAqS samples (on average 3.7). Similar trends were also found using the Simpson index (Table 4). The highest number of observed individual OTUs and species diversity was obtained from GAL, which can be related to the total number of reads sequenced (Table 3).

Table 4 Statistical measures of diversities

Site	**Observed OTUs**	**Chao1**	**Shannon's index**	**Simpson's index**
DAL	58	58	4.06	0.98
GAL	70	70	4.25	0.99
BLA	56	56	4.03	0.98
LADP5	37	37	3.61	0.97
LADP6	45	45	3.81	0.98
LA	35	35	3.56	0.97
MUP	48	48	3.87	0.98

The Bray- Curtis and Jaccard dissimilarity values of DAL vs GAL, DAL vs BLA and GAL vs BLA are all less than 0.67 indicating relatively more OTUs shared among themselves than with either the lake samples or MUP (Table 5). The smallest dissimilarity values were observed among the sample sites of Lake As’ale. The minimum spanning tree for evaluating group separation categorized the sampling sites based on MetaPhlAn2 taxonomic profiling into two distinct groups (Fig 7). Accordingly, DAL, GAL and BLA were clustered in one group while the three coastal sites of Lake As’ale and MUP were grouped in another. This visual categorization of the sampling sites also confirmed that the ecological distance between DAL and BLA is minimum compared to GAL in one group while MUP is out grouped in the other.



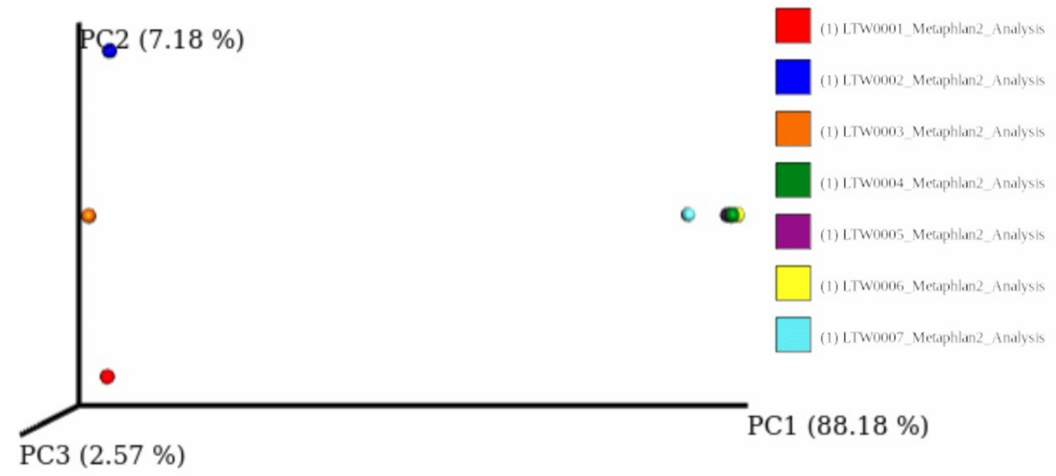
Tree was generated by calculating the Bray-Curtis distances between the samples based on MetaPhlAn2 taxonomic profiling from direct metagenomic shotgun sequencing data using R studio. LTW0001_nt_seq (DAL), LTW0002_nt_seq (GAL), LTW0003_nt_seq (BLA), LTW0004_nt_seq (LADP5), LTW0005_nt_seq (LADP6), LTW0006_nt_seq (LA), LTW0007_nt_seq [MUP]

Table 5 Statistical distance measures

First sample	**Second sample**	**No. observed species in 1st sample**	**No. observed species in 2nd sample**	**No. observed shared species**	**Bray-Curtis dissimilarity**	**Jaccard classic**
DAL	GAL	58	70	32	0.50	0.67
DAL	BLA	58	56	31	0.46	0.63
DAL	LADP5	58	37	5	0.89	0.94
DAL	LADP6	58	45	4	0.92	0.96
DAL	LA	58	35	2	0.96	0.98
DAL	MUP	58	48	12	0.77	0.87
GAL	BLA	70	56	32	0.49	0.66
GAL	LADP5	70	37	7	0.87	0.93
GAL	LADP6	70	45	4	0.93	0.96
GAL	LA	70	35	5	0.90	0.95
GAL	MUP	70	48	13	0.78	0.88
BLA	LADP5	56	37	3	0.94	0.97
BLA	LADP6	56	45	2	0.96	0.98
BLA	LA	56	35	2	0.96	0.98
BLA	MUP	56	48	10	0.81	0.89
LADP5	LADP6	37	45	32	0.22	0.36
LADP5	LA	37	35	31	0.14	0.24
LADP5	MUP	37	48	33	0.22	0.37
LADP6	LA	45	35	33	0.18	0.30
LADP6	MUP	45	48	36	0.23	0.37

LA	MUP	35	48	32	0.23	0.37

The PcoA was used to assist with subjective classifications and visualize the level of similarity of individual microbial compositions of sample sites by detecting pattern. The Bray-Curtis dissimilarity matrices computed for the non assembled metagenome taxonomy profiles, captured more than 80% of the variation in the first axis of the ordination plot (Fig 8). The three shorter distances in the ordination plots were obtained for LA, LADP5 and LADP6 indicating high similarity while the longer distance was for GAL and MUP indicating high dissimilarity.



PcoA was calculated on BIOM- formatted taxonomic profiles using a Qiime utility script for MetaPhlAn2 taxonomy profile where the first principal coordinate contained 88.18% and the second had 7.18%. LTW0001_Metaphlan2_Analysis (DAL); LTW0002_Metaphlan2_Analysis (GAL); LTW0003_Metaphlan2_Analysis (BLA); LTW0004_Metaphlan2_Analysis (LADP5); LTW0005_Metaphlan2_Analysis (LAPD6); LTW0006_Metaphlan2_Analysis (LA); LTW0007_Metaphlan2_Analysis (MUP).

3.3.4. Comparison of microbial communities

The resulted OTU profile displayed the absence of direct mapping of sequence reads to eukaryotes specific marker genes (Fig. 9A). Bacteria dominated the sequence reads from DAL, GAL and BLA while sequence reads mapped to archaea predominate LA, LADP5, LADP6 and MUP samples (Fig. 9A and Fig. 10). However, for BLA metagenome, MetaPhlAn 2 could not map sequence reads to any archaeal marker genes. Total of 155 different OTUs from 6 prokaryotic phyla and one viral division were represented in all studied sites (Fig. 9B and Fig. 10). In coastal sites of Lake As'ale and MUP samples, the main phyla retrieved were Euryarchaeota, where an average of 82% and 69% of total OTUs were grouped respectively. In DAL, GAL and BLA samples, the main phyla retrieved from the bacterial OTUs were Proteobacteria (50%, 44% and 51%, respectively), Firmicutes (22%, 17.5% and 21% respectively), Actinobacteria (15%, 14% and 19%, respectively) and Verrucomicrobia (1.1%). In GAL Bacteroidetes was the second abundant phylum, which accounted 19% of the total OTUs profiled (Fig. 9B and Fig. 10).

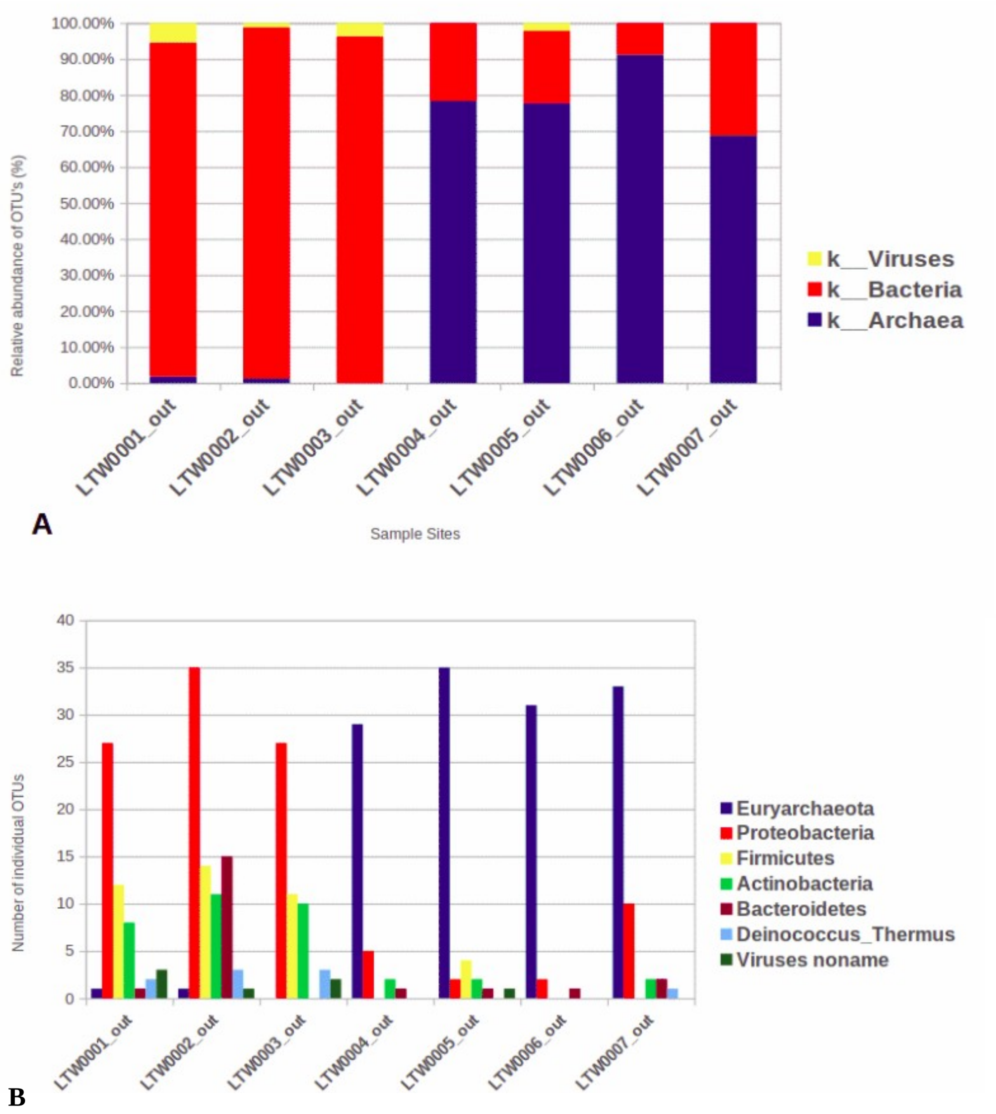


Fig. 9 Domain-level (A) and Phylum-level (B) distribution of microbes

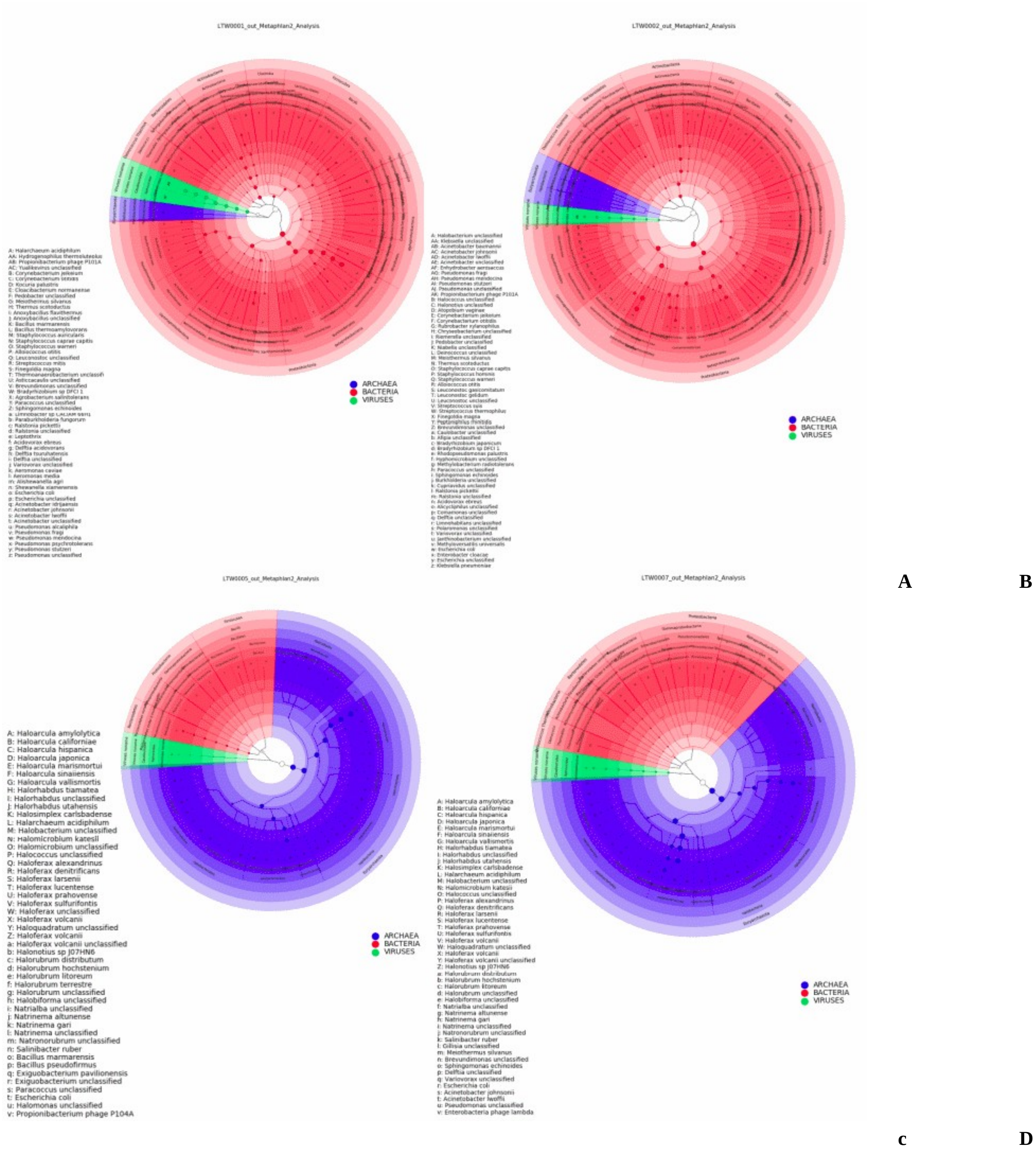

LTW0001_out_Metaphlan2_Analysis
A: Halarchaeum acidiphilum
AA: Hydrogenophilus thermoluteolus
AB: Propionibacterium phage P101A
AC: Yualikevirus unclassified
B: Corynebacterium jeikeium
C: Corynebacterium striatum
D: Kocuria palustris
E: Cloacibacterium normanense
F: Pedobacter unclassified
G: Meiothermus silvanus
H: Thermus scotoductus
I: Anoxybacillus flavithermus
J: Anoxybacillus unclassified
K: Bacillus marmarensis
L: Bacillus thermoamylovorans
M: Staphylococcus auricularis
N: Staphylococcus caprae capitis
O: Staphylococcus warneri
P: Alloiococcus otitis
Q: Leuconostoc unclassified
R: Streptococcus mitis
S: Finegoldia magna
T: Thermoanaerobacterium unclassifi
U: Anticcacaulis unclassified
V: Brevundimonas unclassified
W: Bradyrhizobium sp DFCI 1
X: Agrobacterium salinitolerans
Y: Paracoccus unclassified
Z: Sphingomonas echinoides
a: Limnobacter sp CACIAM 66H1
b: Paraburkholderia fungorum
c: Ralstonia pickettii
d: Ralstonia unclassified
e: Leptothrix
f: Acidovorax ebreus
g: Delftia acidovorans
h: Delftia tsuruhatensis
i: Delftia unclassified
j: Variovorax unclassified
k: Aeromonas caviae
l: Aeromonas media
m: Alishewanella agri
n: Shewanella xiamenensis
o: Escherichia coli
p: Escherichia unclassified
q: Acinetobacter idrijaensis
r: Acinetobacter johnsonii
s: Acinetobacter lwoffii
t: Acinetobacter unclassified
u: Pseudomonas alcaliphila
v: Pseudomonas fragi
w: Pseudomonas mendocina
x: Pseudomonas psychrotolerans
y: Pseudomonas stutzeri
z: Pseudomonas unclassified
● ARCHAEA
● BACTERIA
● VIRUSES
LTW0002_out_Metaphlan2_Analysis
A: Halobacterium unclassified
AA: Klebsiella unclassified
AB: Acinetobacter baumannii
AC: Acinetobacter johnsonii
AD: Acinetobacter lwoffii
AE: Acinetobacter unclassified
AF: Enhydrobacter aerosaccus
AG: Pseudomonas fragi
AH: Pseudomonas mendocina
AI: Pseudomonas stutzeri
AJ: Pseudomonas unclassified
AK: Propionibacterium phage P101A
B: Halococcus unclassified
C: Halonotius unclassified
D: Atopobium vaginae
E: Corynebacterium jeikeium
F: Corynebacterium otitidis
G: Rubrobacter xylanophilus
H: Chryseobacterium unclassified
I: Riemerella unclassified
J: Pedobacter unclassified
K: Niabella unclassified
L: Deinococcus unclassified
M: Meiothermus silvanus
N: Thermus scotoductus
O: Staphylococcus caprae capitis
P: Staphylococcus hominis
Q: Staphylococcus warneri
R: Alloiococcus otitis
S: Leuconostoc gasicomitatum
T: Leuconostoc gelidum
U: Leuconostoc unclassified
V: Streptococcus suis
W: Streptococcus thermophilus
X: Finegoldia magna
Y: Peptoniphilus rhinitidis
Z: Brevundimonas unclassified
a: Caulobacter unclassified
b: Afipia unclassified
c: Bradyrhizobium japonicum
d: Bradyrhizobium sp DFCI 1
e: Rhodopseudomonas palustris
f: Hyphomicrobium unclassified
g: Methylobacterium radiotolerans
h: Paracoccus unclassified
i: Sphingomonas echinoides
j: Burkholderia unclassified
k: Cupriavidus unclassified
l: Ralstonia pickettii
m: Ralstonia unclassified
n: Acidovorax ebreus
o: Alicycliphilus unclassified
p: Comamonas unclassified
q: Delftia unclassified
r: Limnohabitans unclassified
s: Polaromonas unclassified
t: Variovorax unclassified
u: Janthinobacterium unclassified
v: Methyloversatilis universalis
w: Escherichia coli
x: Enterobacter cloacae
y: Escherichia unclassified
z: Klebsiella pneumoniae
● ARCHAEA
● BACTERIA
● VIRUSES
A
B
LTW0005_out_Metaphlan2_Analysis
A: Haloarcula amylolytica
B: Haloarcula californiae
C: Haloarcula hispanica
D: Haloarcula japonica
E: Haloarcula marismortui
F: Haloarcula sinaiiensis
G: Haloarcula vallismortis
H: Halorhabdus tiamatea
I: Halorhabdus unclassified
J: Halorhabdus utahensis
K: Halosimplex carlsbadense
L: Halarchaeum acidiphilum
M: Halobacterium unclassified
N: Halomicrobium katesii
O: Halomicrobium unclassified
P: Halococcus unclassified
Q: Haloferax alexandrinus
R: Haloferax denitrificans
S: Haloferax larsenii
T: Haloferax lucentense
U: Haloferax prahovense
V: Haloferax sulfurifontis
W: Haloferax unclassified
X: Haloferax volcanii
Y: Haloquadratum unclassified
Z: Haloferax volcanii
a: Haloferax volcanii unclassified
b: Halonotius sp J07HN6
c: Halorubrum distributum
d: Halorubrum hochstenium
e: Halorubrum litoreum
f: Halorubrum terrestre
g: Halorubrum unclassified
h: Halobiforma unclassified
i: Natrialba unclassified
j: Natrinema altunense
k: Natrinema gari
l: Natrinema unclassified
m: Natronorubrum unclassified
n: Salinibacter ruber
o: Bacillus marmarensis
p: Bacillus pseudofirmus
q: Exiguobacterium pavilionensis
r: Exiguobacterium unclassified
s: Paracoccus unclassified
t: Escherichia coli
u: Halomonas unclassified
v: Propionibacterium phage P104A
● ARCHAEA
● BACTERIA
● VIRUSES
LTW0007_out_Metaphlan2_Analysis
A: Haloarcula amylolytica
B: Haloarcula californiae
C: Haloarcula hispanica
D: Haloarcula japonica
E: Haloarcula marismortui
F: Haloarcula sinaiiensis
G: Haloarcula vallismortis
H: Halorhabdus tiamatea
I: Halorhabdus unclassified
J: Halorhabdus utahensis
K: Halosimplex carlsbadense
L: Halarchaeum acidiphilum
M: Halobacterium unclassified
N: Halomicrobium katesii
O: Halococcus unclassified
P: Haloferax alexandrinus
Q: Haloferax denitrificans
R: Haloferax larsenii
S: Haloferax lucentense
T: Haloferax prahovense
U: Haloferax sulfurifontis
V: Haloferax volcanii
W: Haloquadratum unclassified
X: Haloferax volcanii
Y: Haloferax volcanii unclassified
Z: Halonotius sp J07HN6
a: Halorubrum distributum
b: Halorubrum hochstenium
c: Halorubrum litoreum
d: Halorubrum unclassified
e: Halobiforma unclassified
f: Natrialba unclassified
g: Natrinema altunense
h: Natrinema gari
i: Natrinema unclassified
j: Natronorubrum unclassified
k: Salinibacter ruber
l: Gillisia unclassified
m: Meiothermus silvanus
n: Brevundimonas unclassified
o: Sphingomonas echinoides
p: Delftia unclassified
q: Variovorax unclassified
r: Escherichia coli
s: Acinetobacter johnsonii
t: Acinetobacter lwoffii
u: Pseudomonas unclassified
v: Enterobacteria phage lambda
● ARCHAEA
● BACTERIA
● VIRUSES
c
D

Fig. 10 Phylogentic tree produced by GraPhlAn

Species-level OTUs from MetaPhlAn2 taxonomy profile. DAL(**A**), GAL (**B**), BLA (**E**), LADP5 (**F**), LADP6 (**C**), LA (**G**) and MUP (**D**). Name of OTUs are listed in Appendix 2.

Five archaeal families of class Halobacteria, namely Halobacteriaceae, Haloarculaceae, Halococcaceae, Natriabaceae and Halorubraceaea, were identified in HsAqS sample sites. From these archaeal families, the most abundantly profiled were Halorubraceae (total average of 74.5%) and Halobacteriaceae (68.2%) in Lake As’ale and MUP, respectively.

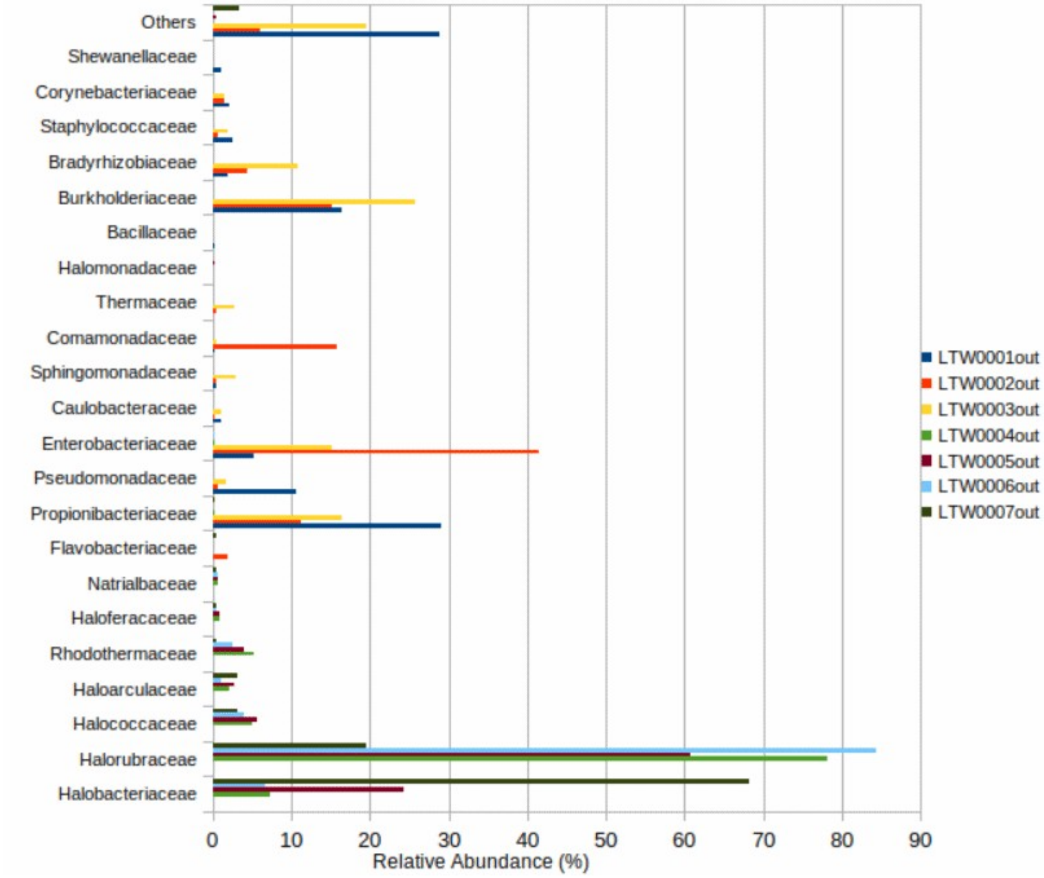


Fig. 11 Top 22 abundant families

List of top 22 abundant families profiled in all studied sample sites using MetaPhlAn 2. LTW0001_out (DAL), LTW0002_out (GAL), LTW0003_out (BLA), LTW0004_out (LADP5), LTW0005_out (LADP6), LTW0006_out (LA), LTW0007_out (MUP).

For HsAP samples, only Halobacteriaceae (less than 0.05%) was profiled in DAL while Halobacteriaceae (0.03%), Halorubraceae (0.03%)and Halococcaceae (less than 0.02%) were profiled in GAL (Fig. 11). The bacterial families Propionibacteriaceae, Burkholderiaceae, Enterobacteriaceae and Bradyrhizobiaceae, are among the top ten abundantly and commonly profiled in the three HsAP sample sites. However, while Pseudomonadaceae (10.6%) and Comamonadaceae (15.8%) were identified as abundant families in DAL and GAL respectively, they were scarce in BLA (1.6% and 0.5% respectively). On the contrary, the total percentage of bacterial OTU is less than 6% in HsAqS and the Rhodothermaceae (average of 3.9% in Lake As'ale coastal sites and 0.5% MUP) was relatively abundant than other bacterial families. Yet this bacterial family was absent in all three HsAP sample sites (Fig. 11).

Unclassified OTUs at different levels of taxonomic ranks were obtained from the MetaPhlAn 2 analysis. Large amount (up to 79%) in DAL, GAL and BLA was not identified at genus and species levels (Table 6). On the other hand, for the three coastal sites of Lake As'ale, an average of 43% of the total profiled OTUs were not identified at genus and species levels (Table 6 and Appendix 2). For MUP, more than 48% of the profiled OTUs were not identified at genus and species levels. In DAL, GAL, BLA, LADP5 and MUP, small percentage of OTUs (0.05%, 0.07%, 0.07%, 0.02% and 0.04%, respectively) were not identified order and family taxonomic levels (Table 6).

Table 6 Percentage of unclassified OTUs based on non assembled sequence reads

	DAL	**GAL**	**BLA**	**LADP5**	**LADP6**	**LA**	**MUP**
Total no. of OTUs	58	70	56	37	45	35	48
Total no. of unclassified OTUs at:							
Order and Family	3	5	4	1	0	0	2
Genera	14	28	13	11	13	11	14
Species	29	28	29	6	5	4	7
Percentage of unclassified OTUs (%)	79.3	87.1	82.1	48.6	40	42.9	48
Relative abundance of sequence reads mapped for unclassified hits (%)	80.4	85.2	72.9	20	37.7	18.7	76.2

The heatmap generated by MetaPhlAn2 depicting the top 25 abundantly presented OTUs is shown in Fig 12. In DAL and BLA, unclassified strain of *Paraburkholderia fungorum* was the most abundant profiled OTU constituting more than 30% of the reads (Fig. 13B). For GAL, unclassified *Escherichia* was highly abundant followed by unclassified genera *Burkholderia* (more than 30% and 12% of the reads respectively).

In the case of the coastal sites of Lake As’ale, the top three highly abundant OTUs are *Halonotius sp. J07HN6 GCF_000416025, unclassified Halobacterium* and unclassified strain of *Salinibacter ruber,* respectively (Fig. 13A). However, for MUP, the top tree profiled OTUs are unclassified *Halobacterium* (50.7%), *Halonotius sp.*

J07HN6 GCF_000416025 (18.2%) and unclassified strain of *Halarchaeum acidiphilum* (17.5%) respectively (Fig. 13A). The only identified OTU commonly shared among all sampling sites was unclassified *Escherichia coli* (Appendix 2).

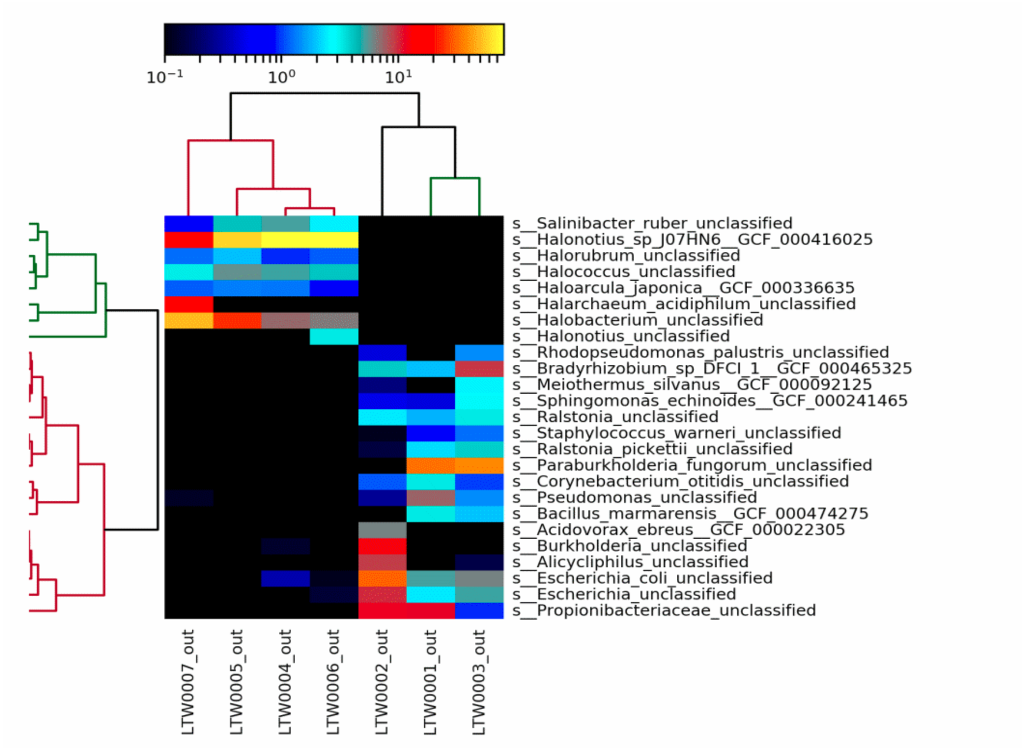


Fig. 12 Heatmap of top 25 abundant OTUs generated by MetPhlAn2

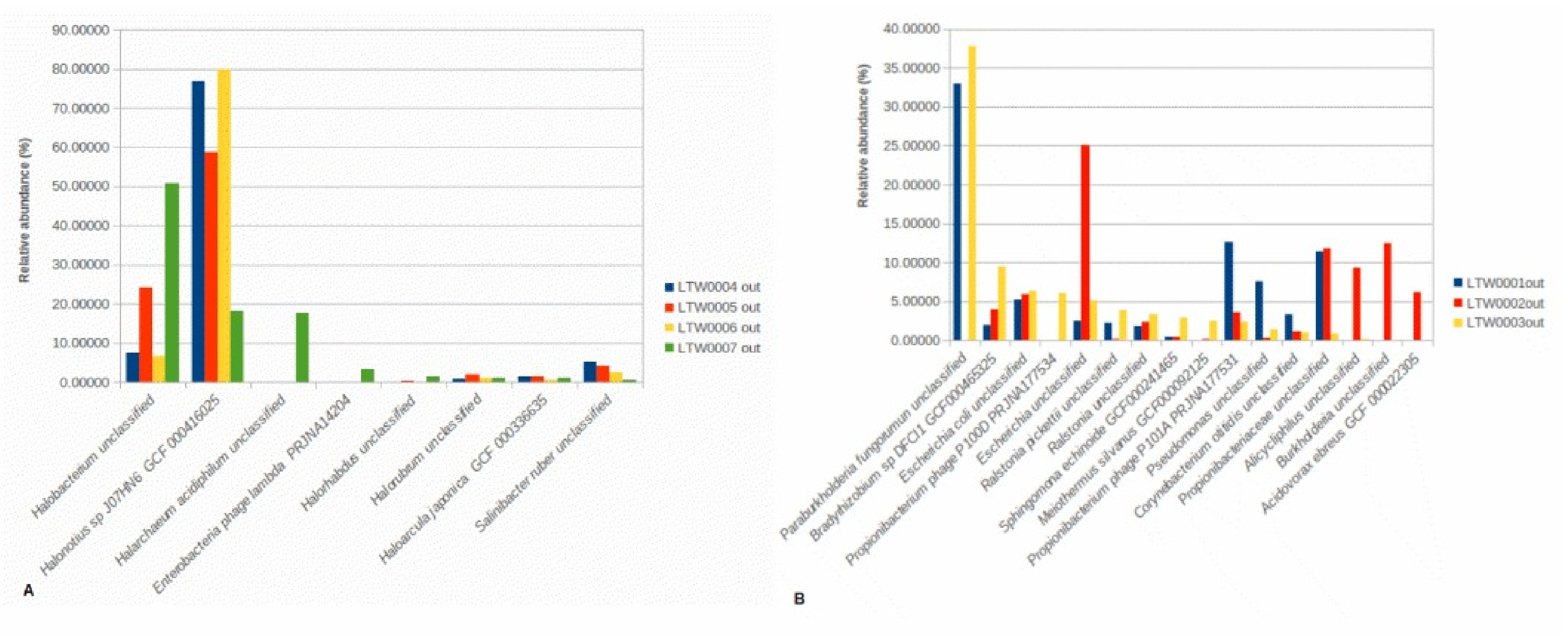


Fig. 13 Bar graph of highly abundant OTUs

The graph shows the top 8 abundant OTUs of HsAqS (**A**) and top 10 OTUs of HsAP (**B**) that constituted more than 90% and 80% of the abundance respectively. LTW0001_out (DAL), LTW0002_out (GAL), LTW0003_out (BLA), LTW0004_out (LADP5), LTW0005_out (LADP6), LTW0006_out (LA), LTW0007_out (MUP).

3.4. Discussion

3.4.1. Geochemical and physico-chemical properties

The salinity and pH measurements of the ponds and the lake results listed in Table 1. Solfataric fields are poly-extreme environments found all over the world in North America (e.g. Yellowstone National Park), Europe (e.g. Campi Flegrei), Asia (e.g. Hokkaido), Africa (e.g. Dallol) and so on and are characterized by low pH and high temperatures (Mayer *et al.,* 2017; Crognale *et al.,* 2018). The year 2015 was special for investigating the extreme hydro-geothermal environments of Danakil Depression as a result of phreatic eruption at the Dallol summit crater, many of the previously existing acid ponds dried up and few new ones appeared (Asrat, 2016). Thence, as as consequence, the size and volume of the three studied HsAP were affected by the latest periodic phreatic eruption (Nobile *et al.,* 2010; Franzson *et al.,* 2015; Asrat, 2016).

The five studied locations in the Danakil Depression were grouped into two extreme environment categories: Hypersaline Acidic Ponds [HsAP] (DAL, GAL and BLA) and Hypersaline Aquatic Systems [HsAqS] (Lake As’ale and MUP) (Thiel, 2011; Gomez, 2014). The liquid brine samples of the HsAP were composed of fluids derived mainly from a mixture of groundwater of meteoric origin and hydrothermal-geothermal up-flow (Franzson *et al.,* 2015; Asrat, 2016; Yeh and Lee, 2018). At the time of sampling in 2015, the isotopes $\delta^{18}O$ and $\delta^{2}H$ measured showed that the sources of water for GAL and BLA were different and the DAL brines had strong volcanic signature while GAL and BLA were fed by non-volcanic, dominantly rainfall originated groundwater. In addition, the three HsAP were different with respect to the total percent salt concentration, where the total salinity of GAL and BLA were the highest (68% and 60% respectively). Such kind of super saturation of salts can only be achieved under very acidic condition, which was a fact in the case of the three acid pools according to the pH result.

The results from measurement of ions concentration conferred important information on the dynamic conditions in the Danakil Depression. For example, in the current study, the highest record of Mg^{+2} concentration was in BLA (~12M) with an almost 1:1 ratio to Cl^{-}. This anomalous concentration of Mg^{+2}, potentially was related to the bischofite (hydrous $MgCl_2$) eruption in close proximity to BLA in 2014, which can also be considered inhibitory to many cellular systems (Hallsworth *et al.,* 2007). Under normal conditions, $MgCl_2$ is a stable brine formed by an irreversible reaction of $Mg(s)$ + diluted $2HCl(aq) \rightarrow H_2(g) + MgCl_2(aq)$. However, the extreme environmental conditions in BLA (the very low pH and intense hydrothermal activities due to the ever present volcanic activities in Danakil Depression), can improve water balance and alter the stability of $MgCl_2$ respectively (Herbstein *et al.,* 1982; de Bakker *et al.*, 2011). For Example, at higher temperature (>560 ^{0}C), breakdown of $MgCl_2 \cdot 6H_2O$ to MgO occurs with, MgOHCl as an intermediate decomposition product for the reaction $MgCl_2 \cdot 6H_2 \rightarrow MgOHCl(s) \rightarrow MgO(s) + Hcl(g)$ (Herbstein *et al.*, 1982; de Bakker *et al.,* 2011).

Therefore, upon our analysis of the chemistry of the brine, high concentrations of Mg_2^+ and Cl^- together with large amount of oxidizable organic materials (high COD) and sufficient amount of NO_2^- and NO_3^- indicative of nitrogen cycle were obtained. High quantity of COD can be directly linked to the extreme concentration of hydrothermal gases and fluids (water vapour, CO_2, CH_4, H_2S, ethane, benzene and so on), which are typical to solfatara floors, geothermal hot springs and soils after condensation of steam at shallow depth subsequent to phreatic eruption (Mayer *et al.,* 2017; Crognale *et al.*, 2018)

Contrarily to the three HsAPs, the two HsAqS sampling sites in the Danakil Depression share similarities in their physico-chemical parameters such as salinity, pH and some of the measured ions. The values of oxygen and hydrogen isotopes showed that the northern coastal part of Lake As’ale is characterized by intense evaporation which selectively enrich the coastal water with heavier isotopes. On the other hand, the cold spring in the MUP is less evaporated, mixed and is continuously fed by groundwater from the highlands whose source is predominantly rain fall depleted in $\delta^{18}O$ and δ^2H. Extreme evaporation and insignificant precipitation in the Danakil

Depression is the potential reason for higher concentration of Mg^{2+} in the lake brine (Ahrens, 2011; Fazzini *et al.*, 2015). The outcomes of the physico-chemical analysis put Lake As'ale and MUP in the hyper saline and slightly acidic category of extreme environments lists (Gomez, 2014).

One of the limitations of the study of Lake As'ale is that all the sampling areas of the lake are from the coastal parts, hence it is difficult to infer the holistic biotic and abiotic characteristics of the lake. Nevertheless, in comparison to other slightly acidic hypersaline lakes such as the Aran-Bidgol salt Lake of Iran (Makhdoumi-Kakhki *et al.*, 2012), Lake Tyrell of Australia (Mernagh, 2014) and the Dead Sea (Oren *et al.*, 1995; Oren, 2006; Jacob *et al.*, 2017) the concentrations of Mg^{2+} and Cl^- ions in the coastal brine sample of Lake As'ale are very high. However, the concentration of Mg^{2+} in MUP is similar to some sites of Aran-Bidgol salt Lake and Lake Tyrell, whereas the concentration of Cl^- is still the highest. Yet the amount of SO_4^{2-} is the lowest in both Lake As'ale and MUP as in the case of Dead Sea, which was considered to be one of the hypersaline lakes on earth to have the lowest concentration of SO_4^{2-} (Oren, 2006; Oren, 2010).

3.4.2. Metagenomics, richness and diversity

Cycles of environmental and geological changes in natural systems over different periods of time and amplitudes can affect the microbial composition, genetic repertoire as well as activity of microbial community members, and community function as a whole (Ghebrezgabher *et al.,* 2016). The possibility of existing biota, especially in the HsAP of Danakil Depression have been both intriguing and debatable topic. The difficulty of studying extremophiles using culture-based methods has been well recognized. Though molecular methods have shown supremacy over culture based diversity study of extremophiles, DNA extraction techniques and sequencing strategies play significant role in defining environmental microbial communities (Roux *et al.,* 2011; Poretsky *et al.,* 2014; Ranjan *et al.,* 2015; Bag *et al.,* 2016; Ketchum *et al.,* 2018).

The major challenge of metagenome study is optimizing proper lysis of heterogeneous microbial communities without damaging their genomes. Combination of physical, chemical and mechanical lysis methods for proper lysis of microbial inhabitants and extraction of community DNA from different environments have been proven suitable for different downstream analysis (Bag *et al.,* 2016). In this study, combined physical and chemical lysis techniques for heterogeneous microbial community cells and whole community metagenome sequencing were effectively used to capture taxa that are traditionally difficult to lyse and excessively rare to detect respectively. Usually, environments with extreme conditions are either tolerated or required by very few highly specialized organisms. The consequences of extreme salinity, pH and some ions such as Mg^{2+} are known to impact microbial community compositions and richness in extreme environments.

Generally, diversity and richness of inhabiting microbial communities are influenced by physicochemical characteristics than geographic distances, especially in hypersaline environments located at close proximity (Baas-Becking, 1931; Oren, 2006; Hallsworth *et al.*, 2007; Lanzén *et al.*, 2013; Logares *et al.*, 2013; Podell *et al.*, 2014; Simachew *et al.*, 2016; Yang *et al.*, 2016; Plominsky *et al.*, 2018; Uribe-Lorío *et al.*, 2019; Banda *et al.*, 2020). The sequencing data analysis clearly showed the potential influence of the physicochemical characteristics on the structure and composition of the mixed microbial communities. A marked distinction in the microbial composition from the HsAP communities between sites highly affected by hydrothermal fluids (DAL) and non volcanic ground water (GAL and BLA) was observed. Concomitant distinctive structure of the microbial community were also observed in Lake As’ale and MUP, characterized by lower salinity (< 36%) and slightly higher pH values (pH 4-6).

Hence, highly diverse and equally distributed OTUs were encountered in the studied samples sites. Though a typical Shannon Wiener diversity index ranges from 1.5 to 3.5 and rarely reaches 4.5, the results of the calculated index in this study are all above 3.5 indicating that there is high diversity with in each sampled communities (Gaines *et al.*, 1999; Martin *et al.*, 2016). Furthermore, results from Simpson’s index, which takes into account the number of species present, as well as the relative abundance of each species, indicated that a high diversity in the sampling sites (Gaines *et al.,* 1999; Martin *et al.,* 2016). Richness and diversity of microbes inhabiting the studied extreme environments in Danakil Depression, are influenced by the same physico-chemical factors that affect the microbial communities of the Dead sea and the acid hot springs in Yellow Stone volcanic areas (Nissenbaum, 1975; Oren, 2005; Hallsworth *et al.,* 2007; Oren, 2010).

In this investigation, the types of microbial communities in the studied HsAP and HsAqS of Danakil Depression were identified by mapping direct short eDNA sequence reads with MetaPhlAn 2.0. The result revealed that all sampling sites are inhabited by prokaryotes with no eukaryotic OTU detection. Difference was observed in the types of organisms present in samples of HsAP and HsAqS as indicated by the calculated dissimilariy indices. The coastal sites of Lake As'ale and MUP were dominantly inhabited by archaeal OTUs while DAL, BLA and GAL were dominantly inhabited by bacterial OTUs. This findings correlate with other studies conducted on similarly extreme environments (Nissenbaum, 1975; Johnson *et al.,* 2001; Spear *et al.,* 2006; Oren, 2010; Makhdoumi-Kakhki *et al.,* 2012; Mesa *et al.,* 2017; Schuler *et al.,* 2017). For example, investigation of microbial diversity in hyderothermal areas of Yellowstone, USA and Acid Mine Drainage of Los Rueldos, Spain were mainly affiliated with Proteobacteria (Mesa *et al.,* 2017; Schuler *et al.,* 2017).

3.4.3. Extremophiles of the HsAP of Danakil Depression

Acidophiles are known to be widely distributed in Domains of Bacteria and Archaea (Hugenholtz *et al.,* 1998; Johnson *et al.,* 2001; Johnson and Hallberg, 2003; Mendez-Garcia *et al.,* 2015). However, the bioinformatics outcome in this study from HsAP metagenome data showed that most of the reads generated were from the bacterial domain with little to no reads for archaea. Hydrothermal gases and fluids create acidic conditions and significant thermal variances on the geothermal hot springs and soils, which influence the composition of microbial communities (Crognale *et al.,* 2018). A comparison of the microbial communities at phylum level in these poly-extreme habitats revealed that the composition of the dominant groups were largely similar with each other, with few discernible distinction. The microbial communities in DAL, GAL, and BLA were dominated by Proteobacteria, Firmicutes and Actinobacteria, except for GAL which showed a composition of major group Bacteroidetes. These phyla are commonly known for their dominance in acidic hot springs all over the world (Song *et al.,* 2013; Badhai *et al.,* 2015).

However, recently published studies regarding life in the acid ponds and hot springs of Danakil Depression based on 16S rDNA investigation reported contrasting results (Belilla *et al.,* 2019; Gómez *et al.,* 2019). While Gómez *et al.*, (2019) confirmed the existence of Nanohaloarchaea in the salt precipitates at DAL hydrothermal fluid source, Belilla *et al.,* (2019) strongly debated the presence of active life forms in DAL and BLA. However, the previous exploration for signs so life in Dallol sulfur springs, using patterns of lipid biomarkers and stable isotope composition was successful enough to identify microbial community largely composed of thermophilic members of the Aquificae, Thermotogae, Chloroflexi and Proteobacteria phyla (Carrizo *et al.,* 2019).

Distinct communities of microbiota was observed in DAL, GAL and BLA, even if Proteobacteria is the dominant phyla in all. This observed taxonomic pattern is similar to the acid pools of Yellowstone but different in many other acid mine drainage sites studied so far (Hugenholtz *et al.,* 1998; Johnson *et al.,* 2001; Spear *et al.,* 2006). The only archaeal phylum identified in the acid ponds is the Euryarchaeota while Crenarchaeote is the dominant one in previously studied acidic hydrothermal fields such as the Rehai hot springs of China and the Yellowstone hot springs (Spear *et al.,* 2006; Hou *et al.,* 2013).

Among the top ten abundantly found bacterial families in the three HsAP, Propionibacteriaceae, Burkholderiaceae, Enterobacteriaceae, and Bradyrhizobiaceae were commonly profiled. These four families are well known for constituting species of various habitats (Badhai *et al.,* 2015; Scholz and Kilian, 2016; Carrier and Reitzel, 2017). Similar species of Burkholderiacaea have found to be adapting and surviving variety of environments (Weber and King 2017; Jin *et al.,* 2020; Tan *et al.,* 2020). The lack of clear demarcating of closely related species in Burkholderiacaea made differentiation of new strains from similar species in different environments more difficult (Jin *et al.*, 2020; Tan *et al.*, 2020). Furthermore, different types of virus OTUs were also identified, especially from the Siphoviridae family, which are known to infect different species in the genera *Pseudomonas* and *Propionibacterium* (Ackermann and Kropinski, 2007; Brüggemann and Lood, 2013).

Many of the OTUs identified in this study, especially in DAL, BLA and GAL, are unclassified at different levels of taxonomic ranks. Only few of the OTUs were identifiable at a strain level while many remain unclassified. This finding correlates with other extreme environments that were systematically studied for the first time, as newly explored poly-extreme habitats are troves for novel organisms (Stetter, 1982; Hugenholtz *et al.*, 1998; Cavicchioli *et al.*, 2002; van den Burg, 2003). The unclassified strain *Paraburkholderia fungorumun* was the most abundant OTU in DAL and BLA. This Gram-negative microorganism is famous for it’s benefits in agriculture and other sectors as an agent for biocontrol and bioremediation as well as pathogenic characteristics (Kaur *et al.*, 2017; Tan *et al.*, 2020; Vio *et al.*, 2020). In previous studies on *P. fungorumun* strain isolated from *Manis javanica*, different genes related to stress responses such as osmotic stress, detoxification, heat shock and cold shock, were identified (Tan *et al*., 2020). The identification of these stress response genes shed light on how *P. fungorumun* adapt and survive in a variety of environments (Tan *et al*., 2020).

In GAL however, the most abundant OTU was the unclassified *Escherichia*. Though the genus *Escherichia* has been associated with humans as pathogenic bacteria, this group of organisms has the ability to survive on or sub surface under intense cosmic rays for thousands of years (Merino *et al.,* 2019). It is not unusual to find pathogenic microorganisms surviving as a free-living organism in extreme environments such as acidic pools and salt mines (Sheehan *et al*., 2005; Akhtar *et al.*, 2008; van Elsas *et al.*, 2011). Furthermore, environmental strains of *E. coli,* which are the most heterogeneous phylogroups due to the diverse nature of their environmental niche. have been identified from different environments ranging from tropical environments to marine or to wastewater treatments, in the presence or absence of human or animal contamination (Ishii and Sadowsky 2008; Goto and Yan 2011; Abram *et al.*, 2020).

Among all identified OTUs in HsAP, only *Meiothermus silvanus*, *Bradyrhizobium_sp_DFCI_1__GCF_000465325* and *Sphingomonas echinoides* were commonly identified and classified at strain level. Generally, the list of bacterial OTUs profiled in the three HsAp indicates the biotechnological potential of the sites. To mention a few, *Rubrobacter xylanophilus* and *Deinococcus geothermalis* (Ferreira *et al.,* 1997; Ferreira *et al.,* 1999) are bacteria well known for their ability to biodegrade radioactive wastes, while *Methylobacterium radiotoleranst* and *Acidovorax ebreus* are known for removing polycyclic aromatic hydrocarbon wastes and for enhancing oil recovery respectively (Zhu *et al.,* 2013; Nzila *et al.,* 2016). *Meiothermus silvanus* and *Sphingomonas echinoides* are also used in the paper industry as an anti-biofouler and in removing aromatic wastes (Kim *et al.*, 2006; Sikorski *et al*., 2010).

In general, the DAL summit crater, the cold spring of GAL and the satellite eruption of BLA, has an extraterrestrial appearance and quality, which are rightly considered inhospitable to living beings. Yet both molecular based as well as lipid biomarker and carbon stable isotope studies conducted on different DAL hydrothermal systems and solfataric fields signal the presence of life (Carrizo *et al.,* 2019; Gómez *et al.*, 2019). Alternative to the postulation of the starting of life in deep-sea hydrothermal vents, volcanic pools and hot springs on land also provides basic nutrients, energy and conditions as well as reactions to create complex molecules which are precursors of life (Deamer and Damer, 2017, and references therein; Van Kranendonk *et al.* 2017). Hence, the acid ponds scattered in the Danakil Depression are ideal places to study how the poly-extreme environments affect organic chemistry and life on Earth, or on other sites in the solar system (Wächtershäuser, 2006; Barbieri and Cavalazzi.2014).

3.4.4. Extremophiles of the HsAqS of Danakil Depression

Extremely halophilic archaea, belonging to the class Halobacteria of the phylum Euryarchaeota are the dominant OTUs in the coastal parts of Lake As’ale and MUP. Organisms belonging to this group of archaea are well known for being common inhabitants of hypersaline environments (Burns *et al.,* 2004; Yang *et al.,* 2007; Oren, 2010; Sorokin *et al.,* 2014; Quadri *et al.,* 2016). The most common and abundant archaeal OTU in the coastal parts of Lake As’ale and MUP is *Halonotius J07HN6,* which was previously identified as inhabitant of Lake Tyrrell in Australia (Podell *et al.,* 2014). The *Halonotius,* in general, are common inhabitants of hypersaline environments and *Halonotius J07HN6* is believed to be the smallest cellular organism next to *Nanohaloarchaea* (Mernagh *et al.,* 2014; Podell *et al.,* 2014; Durán-Viseras *et al.,* 2019).

The only archaeal OTU that was represented in all coastal sites of Lake As’ale but not in MUP was *Halorubrum terrestre.* This halophilic archaea was first isolated from saline soil in Russia (Ventosa *et al.,* 2004). Phylogentically, the *H. terrestre* is closely related to *H. distributum* and *H. litoreum* (Ventosa *et al*., 2004: Cui *et al.,* 2007). On the contrary, all archaeal OTUs found in MUP were also present at least in one of the sampling sites of the Lake As’ale.

Unclassified genera of *Halorubrum, Halobiforma, Haloquadratum, Natrialba* and *Natronorubrum* were well distributed OTUs in the two HsAqS. As in the case of HsAqS, many of the mapped reads were not classified to genus or species level. Two main reasons can be put forward to justify the results. First the two HsAqS environments are unique poly-extreme environments which have not been biologically studied before, hence might be troves for novel organisms (Hugenholtz *et al.,* 1998; Cavicchioli *et al.,* 2002; van den Burg, 2003). Second, most of the previously generated sequences data related to the genera from our study, were from hypersaline and alkaline environments (Sorokin *et al.,* 2014; Quadri *et al.,* 2016; Simachew *et al.,* 2016) and distantly related phylotypes of these genera were only recovered in acidic environments like the Dead Sea (Oren, 2010). Hence, lack of sufficient but exclusive clade-specific marker genes from slightly acidic hypersaline aquatic environments might affect the calculation of taxonomic classification from phylogenetic distances. Thus, MetaPhlAn 2.0 can be influenced to characterize these organisms as unclassified genera or species (Truong *et al.,* 2015; Teo and Neretti, 2016; Pasolli *et al.,* 2017).

Variation in the number of bacteria OTUs was observed in the coastal sites of Lake As’ale and MUP. Mud’ara Pond had 4 times higher number of bacteria OTUs than in Lake As’ale. Unclassified strain of *Salinibacter ruber* was the abundant and commonly distributed bacterial OTU in coastal parts of Lake As’ale and MUP. This bacterium is one of the common and abundantly found bacterial group together with extremely halophilic members of the Archaea in hypersaline environments and well known for its aerobic and heterotrophic nature (Nissenbaum, 1975; Antón *et al.,* 2002; Mormile *et al.,* 2009; Ventosa *et al*., 2009; González-Torres and Gabaldón. 2018; Viver *et al.,* 2018). Surprisingly, *S. ruber* OTU was not profiled in all samples of HsAP as this bacteria can not survive extreme acidic conditions, but rather prefer neutral to alkaline pH (6.5-8.0) but can grow at pH as low as 5.5 (Antón *et al.,* 2002; Makhdoumi-Kakhki *et al.,* 2012).

Though more sampling is needed from Lake As’ale, for more comprehensive analysis of the lake’s microbial diversity, most of the OTUs identified at species level have been reported in previous studies as biotechnological important. Therefore, the findings of this study on the brine samples can be used as preliminary data for further investigation on the two water bodies of Danakil Depression, namely Lake As’ale and MUP.

3.5. Conclusion

In this investigation, we were able to demonstrate that all sampling sites (i.e., DAL, GAL, BLA, coastal sites of Lake As’ale and MUP) are different extreme habitats found in Danakil Depression. Considering the extremity of the environmental conditions of the Danakil Depression, this metagenomics dataset can be a starting point for studying unique gene functions that enable thriving biocommunities in hypersaline and extremely or moderately acidic conditions. One of the limitations of this study is that all the sample areas of Lake As'ale are from the coastal parts, hence difficulty arise to infer the holistic biotic and abiotic characteristics of the lake. The results of the physico-chemical and the metagenomics analysis showed that there is a discord between these two aquatic environments in contrast to their geographical proximity. Differences in ionic contents such as Mg $^{2+}$ and SO_4^{2-} and pH seemed to affect the diversity and abundance of prokaryotes inhabiting the studied environments. Therefore, the prokaryotes inhabiting the coastal brine of Lake As'ale must be specialized to tolerate the high concentration of divalent ions of, which are characteristics of lakes with low pH. The three acid ponds (DAL, GAL and BLA) are as different in their microbial community composition as they are physicochemically distinct. Each of HsAP has distinct biotia flourishing in them but Proteobacteria is dominant. In general, this preliminary data can be used to get a glance at the microbial diversity and the undeniable biotechnological potential of the Danakil Depression. Despite broad taxonomic similarities among the salt-saturated metagenomes analyzed, ee have identified unique biota associated with each sampling sites. Therefore, further investigation is recommended to understand the individual extremophile's adaptation mechanisms and their role in the ecosystem.

Chapter Four

4. In silico Determination of nutrients and biogeochemical cycles in the extreme environments of Danakil Depression using Functional Metagenomics

Abstract

Predictions of gene function and metabolic pathways based on a categorization system, which organizes gene functional categories into a hierarchy (SEED) and Kyoto Encyclopedia of Genes and Genomes (KEGG) respectively disclosed the presence of different nutrient cycles in the studied HsAP and HsAqS environments. For BLA and DAL, many genes involved in carbohydrate, nitrogen and sulfur cycles were predicted thus metabolic pathway mapping for carbon fixation, nitrogen assimilation and sulfur oxidation. We predict that highly efficient and light-independent carbon fixation takes place in the two HsAqS. On the contrary, genes encoding enzymes involved in hydrogenotrophic and acetoclastic methanogenesis appeared solely in ponds of GAL, implying the biological source of the hazardous methane gas existing in that environment. Special attention was given to DAL’s and

BLA’s partial or completely sequenced genes involved in nitrogen metabolism. The KEGG nitrogen metabolism pathway mapping results for both acid ponds showed that all the predicted genes are involved directly or indirectly in the assimilation of ammonia and no dissimilation or nitrification process was identified. Furthermore, the deduced nitrogen fixation in the two acid ponds based on SEED classifcation indicated the presence of different sets of nitrogen fixing (nif) genes for biosynthesis and maturation of nitrogenase. Based on the investigation of the sources of the genes of interest, it is clear that cooperative interactions between members of the microbial communities and syntrophic metabolism is the main strategy adapted to utilize the micronutrients in all the studied sites. Considering the extremity of the environmental conditions of the studied sites in the Danakil Depression, this metagenomics data set can add value to the study of unique gene functions in nutrient metabolisms that enable thriving biocommunities in polyextreme environmental conditions.

Keywords; Insilico, Methanogenesis, Gaet’ale, Nitrogen fixation, Carbon fixation

4.1. Introduction

Microorganisms are fundamental factors of geochemical cycling and bio-transformation of nutrients in many extreme environments. These extremophilic microorganisms are uniquely adapted to flourish in extreme environments and are particularly important in influencing the global bio-geochemical cycling (Whitman *et* *al.,* 1998; Rothschild and Mancinelli, 2001; Rampelotto, 2013; Selvarajan *et al.,* 2014; Uribe-Loríо *et al.,* 2019). The interdisciplinary researches conducted on extremophiles suggested that no matter how much extreme environments determined the existence of life, the inhabiting living organisms also influence these environments (Kluber *et al.,* 2020; Nayak *et al.*, 2020). These dynamics between the extreme ecosystems and the extremophiles become more apparent when investigating the responses between biological activities and attributes which characterize an ecosystem (Bang *et al.*, 2018; Merino *et al.*, 2019; Kluber *et al.,* 2020; Nayak *et al.*, 2020).

The bio-transformation of minerals and nutrients is highly relevant to understand bio-geochemical cycle in extreme environments. Genomic and physiological adaptations imposed by the extreme conditions on the organisms, are fundamental to understand the interactions between the extreme environments and the organisms (Bang *et al.,* 2018; Merino *et al.*, 2019;). Insights into detailed aspects of the complexities of extreme environments bio-cycle can be catered by identifying the biotransformation of chemical elements using molecular details of microbial pathways coupled with analyses of microbial communities (Wang *et al.,* 2015; Coker, 2016; Osborne *et al.,* 2020).

Fast developments in the field of bioinformatics enabled direct analysis of natural microbial communities from environmental samples and their ecological networks

(Su *et al,* 2014; Wang *et al.,* 2015; Coker, 2016; Tran and Phan, 2020). The integration of in silico analyses to metagenomics allows holistic investigation of extremophiles with biotechnological potential (Smith and Plazas 2011; Barriuso and Martínez 2015; Usman *et al.,* 2017). Synergistically implementing multiple methodologies instead of a single approach can increase the rates of discovering target genes and metabolic pathways from metagenomes (Barriuso and Martínez, 2015; Kodzius and Gojobori, 2015; Alves *et al*, 2017). Hence, cataloging of Open Reading Frames (ORFs) present in the metagenome of extremophilic communities based on metabolic function can be valuable source of information for synthetic biology and genetic engineering to create and/or improve bio-resources (Kodzius and Gojobori, 2015; Li *et al.,* 2016).

In order to explain high-throughput genes and/or proteins using annotation, two approaches (‘gene-by-gene’ and subsystem) have been developed so far. The commonly applied ‘gene-by-gene’ approach focus on annotating the entire genes within a single organism (Koonin and Galperin, 2003). However, this type of genome annotation and analysis lacks specific expertise related to the role of each gene, especially for understanding the metabolism of environmental samples (Overbeek *et al.,* 2005). On the other hand, the subsystems genome annotation approach is centered on a set of functional roles of protein families (subsystems) and on sets of functional roles that tie protein-encoding genes to different subsystems (subsystem connections) instead of individual genes (Overbeek *et al.*, 2005; Overbeek *et al.*, 2014).

Comparison of various statistically significant subsystems and metabolism connections present in metagenome samples can yield new insights into the biology of the extreme environments as well as analyze changes in microbial communities (Overbeek *et al.*, 2005; Rodriguez-Brito *et al*., 2006; Neelakanta and Sultana, 2013; Overbeek *et al.*, 2014; Medeiros *et al.*, 2016; Mendes *et al.*, 2017). Therefore, in silico analyses was conducted to probe the biological network from the metagenome sequeces of the environmental brine samples of the three HsAP, and the two HsAqS of the Danakil Depression soon after the 2015 phreatic eruption. In addition, in order to predict nutrient cycles, all the sampling sites were examined as a self-contained, open system with a distinct functional profile. Further emphasis was given to subsystems uniquely predicted in one or two sample sites that can further characterize the nutrient metabolic pathway mapping. Prediction of structure and modeling of selected proteins was done by retrieving the ORFs of protein families of interest and search for orthology from different protein databases.

4.2. Materials and methods

4.2.1. Assembly and taxonomic assignment of contings

Quality sequence reads were assembled using metaSPADes with a flag ‘meta’ and kmers 21, 33 and 55 (Nurk *et al.,* 2017). The resulted metagenome contigs were aligned against NCBI non-redundant protein database (NCBI-nr) using Double Index Alignment of Next Generation Data (DIAMOND) v0.9.24; Translated Basic Local Alignment Search Tool (BLASTx) with the sensitive mode, frameshift alignment for longer sequences and a default e-value cut-off of 0.001 (Buchfink *et al.,* 2015). The taxonomic assignment of assembled contigs was interactively performed using MEtaGenome analyzer 6 Community Edition (MEGAN6 CE) (Huson *et al.,* 2016; Huson *et al.,* 2018). Since one contig may contain several ORFs, each ORF was considered separately during filtration process where alignments that overlap significantly were grouped into segments, denoting different ORFs. The top segments within the best scoring alignments were taken into account.

Accordingly, MEGAN6 CE assigned the annotated reads onto the NCBI taxonomy tree using settings of Lower Common Ancestor (LCA) algorithm for long read adjusted as follows: min score - 100.0; max expected- 0.01; min percent identity-50; top percent- 10 and LCA coverage 80%. The threshold for minimum support that a taxon requires, as a percentage of assigned reads (min support percent) was adjusted to 0.02 so that taxa that obtained at least 0.02% of all aligned bases are reported. As a principle, adjustment of minimum support percent will increase the ‘level of detection” and improve sensitivity for low-abundance species. After the initial automatic binning step, additional manual inspection was performed. Contigs with uncertain taxonomic association, characterized by mixed blastx hits were moved to the ‘Unassigned’ bin.

4.2.2. Functional gene assignment of assembled contigs

The subsequent assembled and DIAMOND-aligned contigs were annotated against the SEED. The SEED is an annotation environment that supports classification of gene function based on collection of biologically defined subsystems (Overbeek *et al.*, 2005; Huson *et al.*, 2016). The SEED program maps each assembled long read onto a gene that has a known functional role and then into one or more subsystems. MEGAN6 CE computed SEED RefSeq ids to functional roles using the ‘seed2ncbi.gz’ file from the SEED server for BLAST alignment in NCBI-nr database (Huson *et al.,* 2016; Bağcı *et al.,* 2019). Functional gene evaluation and protein identification for nutrient cycles focusing on carbohydrate, nitrogen, phosphorous and sulfur metabolisms were performed.

4.2.3. Pathway mapping of nutrient metabolism

Kyoto Encyclopedia of Genes and Genomes database was used to search and map metabolic pathways in the studied sampling sites (Kanehisa and Sato, 2019). Enzyme Commission (EC) number, or the amino acid sequences of genes for key proteins of the focused nutrient (carbohydrate, nitrogen, phosphorous and sulfur) cycles were retrieved from SEED classification and converted to KEGG Orthology (KO) identifier. The website available "Search and Color Pathway" KEGG mapping tool (https://www.genome.jp/kegg/tool/map_pathway2.html) was used for automatic assignment of KO identifiers to KEGG molecular networks and pathway maps of Energy metabolisms (carbon fixation in photosynthetic organisms, methane metabolism, nitrogen metabolism and sulfur metabolism). Significant maps of energy metabolism with well-defined protein families and pathways were searched from each sampling sites for further comparison and analyses among sampling sites.

4.2.4. Retrieving protein sequences and homology modeling of predicted proteins involved in “nitrogen fixing”

Sampling sites with prediction of complete or near complete “nitrogen cycle” were chosen for further analysis of the “nitrogen fixing” pathway. MEGAN’s ‘Alignment Viewer’ was used to compute and view multiple sequence alignment of contigs from selected sampling sites that have significant matches to reference sequences associated with the different SEED subsystems. The contigs were screened and selected from the most profusely represented OTUs across the sampling sites. During the filtration process, each gene was considered separately, as one contig may contain several protein coding ORFs. Hence alignments that overlap significantly were grouped into segments, denoting different genes. The top best scoring alignments were taken into account. Consensus translated ORFs, mapped and covered, more than 95% of respective reference protein sequence residues were exported in fasta format for further in silico analyses.

4.2.5. Primary and secondary structure analysis of predicted proteins

Primary sequence analysis was performed on the consensus protein sequences using ProtParam tool available at website www.expasy.org/tools/protparam (Gasteiger *et al*, 2005). Sequence similarity of translated ORFs grouped as a nitrogen fixating class was analyzed using NCBI blastp (cut-off value 1e-10) taking as reference Nif protein sequences well characterized from different families. Query translated ORFs with more than 95% coverage of residues of the complete homologous reference's protein sequence were considered as near complete sequences for this study. Only complete translated ORFs were selected for further structural analyses.

Prediction of protein pattern, structural domain, active site and their related function of the selected translated ORFs were performed using InterProscan at website http://wwwdev.ebi.ac.uk/interpro/. In addition, secondary structures were predicted using GOR4, SOPMA available at ExpaSy website http://www.expasy.org/tools using default settings and in accordance with the prediction methods (Geourjon and Delage, 1995; Kouza *et al*, 2017).

4.2.6. 3-Dimention modeling of selected predicted proteins

3-Dimention modeling was performed only for predicted proteins common in both sampling sites that fulfill the above criteria. Hence, common query translated complete or near complete ORFs to the studied samples were submitted to SWISS-MODEL (http://swissmodel.expasy.org) under automatic mode settings (Bienert *et al,* 2017; Waterhouse *et al.*, 2018). The server automatically selects the template from experimentally solved protein structures thereby generating the best possible model according to homology modeling method. The predicted 3-D models were then accessed in the form of PDB files.

4.3. Results

4.3.1. Taxonomic overview of assembled contigs

Individual de novo metagenome assemblies were generated from quality-trimmed sequence reads using MetaSpades. The total contigs generated by MetaSpades with weighted average length for all sampling sites are described in Table 7. The N50 represents the length of the shortest contig for which longer and equal length contigs cover at least 50 % of the assembly. The two highest N50 values obtained were from GAL (2344 bp) and BLA (1126 bp).

The results of the alignments of assembled contigs using DIAMOND are shown in Table 8. From DAL assembled metagenome, more than 80% of the contigs were aligned to NCBI nr database by DIAMOND, while for BLA and MUP, more than 73% and 71% of the contigs, respectively were aligned. The least number of contigs aligned to NCBI nr databas was from GAL (50%), while for the three sites of Lake As’ale an average of 57% of contigs were aligned by DIAMOND. The result of DIAMOND analysis is considered as an estimation of the taxonomical content (“species profile”) of the sample from which the reads were collected and interactively explored by MEGAN6 CE.

Table 7 Overview of Metagenome assemblies with MetaSPAde

Site	Sequence ID	No. of contigs	Max. contig length (bp)	N50 (bp)	Total length (bp)	Metagenome size (bp/ NGS seq, read)
DAL	LTW0001	85336	48588	508	38505818	38505818 / 3480089
GAL	LTW0002	161889	429699	2344	125579784	125579784 / 29741784
BLA	LTW0003	109720	67786	1126	77553274	77553274 / 6147722
LADP5	LTW0004	83343	24552	468	32250632	32250632 / 13103002
LADP6	LTW0005	95819	20601	596	46005719	46005719 / 14021795
LA	LTW0006	98941	27707	483	40523474	40523474 / 10440347
MUP	LTW0007	59342	48514	383	22881057	22881057 / 11016361

Table 8 Taxonomic Binning of assembled reads using DIAMOND and MEGAN6 CE

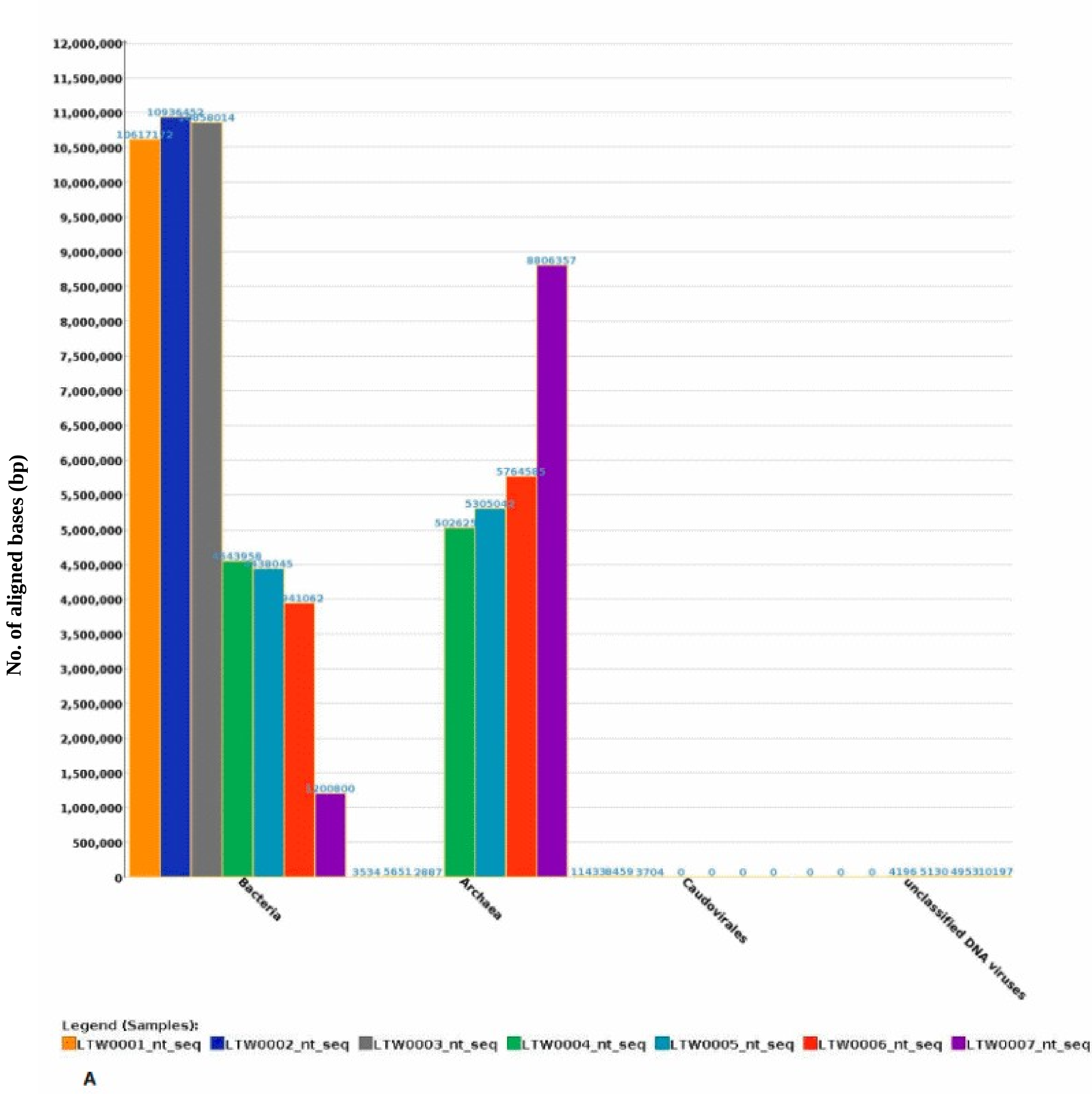
	DAL	**GLA**	**BLA**	**LADP5**	**LADP6**	**LA**	**MUP**
No. of bases that must be assigned to taxon/taxa *	5451	10326	9073	3486	6562	4520	2836
Total No. of contigs aligned by DIAMOND	68991	76755	80595	42305	62841	56019	42549
Average No. of normalized counts of aligned bases per contig assigned by DIAMOND (bases/contig)	154.4	143	135	227.7	156	174.1	236.6
Total No. of normalized counts of aligned bases assigned to NCBI taxonomy (Mb)	10.8	11.1	10.9	9.6	9.8	9.8	10.1
Total No. of normalized count of aligned bases assigned to NCBI taxonomy at species level (Mb)	2.2	3.3	2.7	3.6	3.8	3.7	3.6
Total normalized count with no hit	375	266	240	665	627	637	715
Total normalized count not assigned (Mb)	3.4	3.1	3.2	4.5	4.4	4.4	4.1
No. of OTUs profiled at species level	114	132	134	122	116	138	154
Archaea	0	0	0	83 (70%)	73 (66%)	100 (75%)	128 (84%)
Bacteria	114(100%)	132(100%)	134(100%)	35 (30%)	36 (33.3%)	32 (24.2%)	23 (15%)

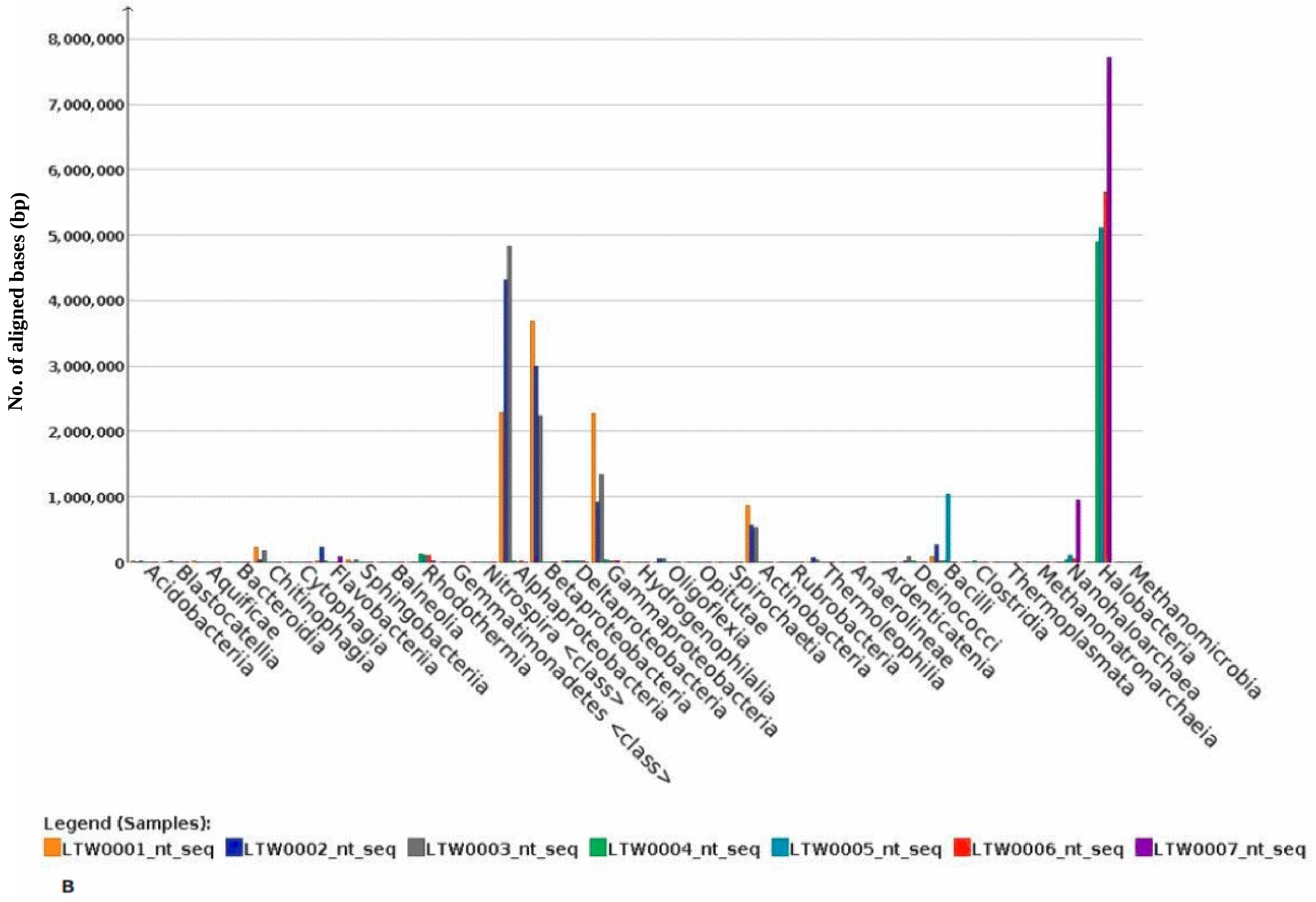
virus	0	0	0	0	1 (0.01%)	1 (0.01%)	2 (0.01%)

(*)=Number of bases that must be assigned to taxon/taxa to appears in the result at Min support percent=0.02; Mb= mega base pair

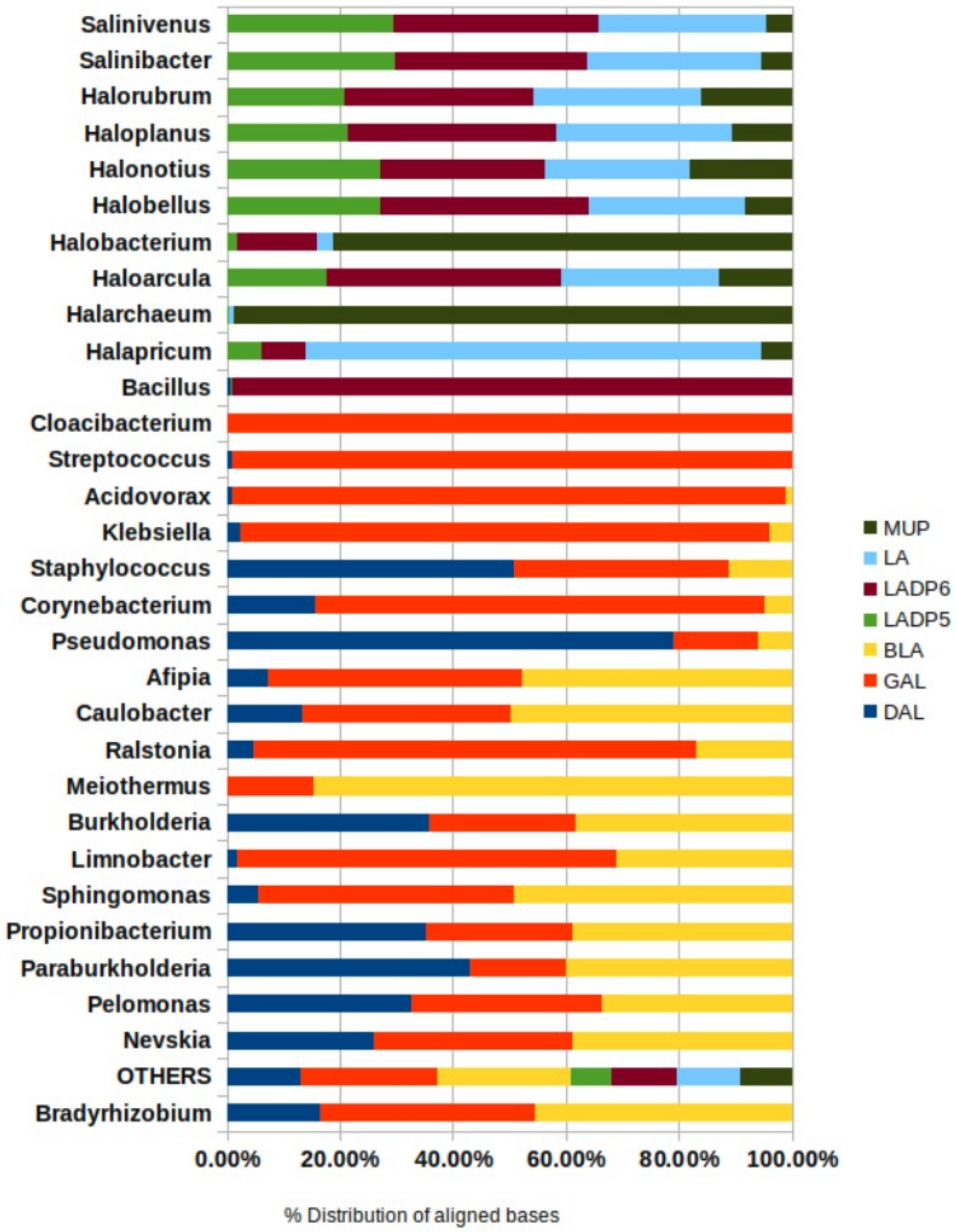
MEGAN6 CE assigned more than 10 million bases from the three HsAP metagenomes to bacteria domain, where more than 7.5 million bases were assigned to the classes Alphaproteobacteria, Betaproteobacteria and Gammaproteobacteria (Fig. 14 A and B). For the metagenomes of the three coastal sites of Lake As’ale, approximate average of 3.9 million and 5.0 million bases were aligned to bacteria and archaea domains, respectively (Fig. 14 A). The highest number of aligned bases for Archaeal domain was obtained from the MUP metagenome with more than 7.5 million bases aligned to class Halobacteria and slightly less than 1 million bases are aligned to Nanohaloarchaea.

In the assembled metagenomes of the three HsAP, more than 86% of the aligned bases parsed to genus level by MEGAN CE were assigned to the top 20 genera, where Bradyrhizobium was the most abundant genus. On the contrary, for the three Lake As’ale’s, Halobellus was the most abundant of the top ten genera that covered an average of 82% of assembled metagenome. In the case of MUP, more than 78% of the aligned bases parsed to genus level were assigned to the top 10 genera, and Halarchaeum is the most abundant genus (Fig 14C). Less than 30% and 40% of aligned bases from HsAP and HsAqS metagenomes were assigned at species level by MEGAN6 CE, respectively (Table 8).


No. of aligned bases (bp)
12,000,000
11,500,000
11,000,000
10,500,000
10,000,000
9,500,000
9,000,000
8,500,000
8,000,000
7,500,000
7,000,000
6,500,000
6,000,000
5,500,000
5,000,000
4,500,000
4,000,000
3,500,000
3,000,000
2,500,000
2,000,000
1,500,000
1,000,000
500,000
0
10617172
10936452
10858014
4543958
4438045
3941062
1200800
3534 5651 2887
502625
5305042
5764585
8806357
114338459 3704
0 0 0 0 0 0 0
4196 5130 495310197
Bacteria
Archaea
Caudovirales
unclassified DNA viruses
Legend (Samples):
LTW0001_nt_seq
LTW0002_nt_seq
LTW0003_nt_seq
LTW0004_nt_seq
LTW0005_nt_seq
LTW0006_nt_seq
LTW0007_nt_seq
A



No. of aligned bases (bp)
8,000,000
7,000,000
6,000,000
5,000,000
4,000,000
3,000,000
2,000,000
1,000,000
0
Acidobacteriia
Blastocatellia
Aquificae
Bacteroidia
Chitinophagia
Cytophagia
Flavobacteriia
Sphingobacteriia
Balneolia
Rhodothermia
Gemmatimonadetes <class>
Nitrospira <class>
Alphaproteobacteria
Betaproteobacteria
Deltaproteobacteria
Gammaproteobacteria
Hydrogenophilalia
Oligoflexia
Opitutae
Spirochaetia
Actinobacteria
Rubrobacteria
Thermoleophilia
Anaerolineae
Ardenticatenia
Deinococci
Bacilli
Clostridia
Thermoplasmata
Methanonatronarchaeia
Nanohaloarchaea
Halobacteria
Methanomicrobia
Legend (Samples):
LTW0001_nt_seq
LTW0002_nt_seq
LTW0003_nt_seq
LTW0004_nt_seq
LTW0005_nt_seq
LTW0006_nt_seq
LTW0007_nt_seq
B



C

Fig. 14 Taxonomic distribution of assembles contigs

Distribution of OTUs at Domain (**A**), Class (**B**) and Genus (**C**) taxonomic level in the sampling sites based on assembled contigs using DIAMOND-MEGAN6CE. LTW0001_nt_seq (DAL), LTW0002_nt_seq (GAL), LTW0003_nt_seq (BLA), LTW0004_nt_seq (LADP5), LTW0005_nt_seq (LADP6), LTW0006_nt_seq (LA), LTW0007_nt_seq (MUP).

4.3.2. Functional classification and annotation of assembled contigs

The absolute total percentage of functionally annotated reads is less than 30% in the assembled metagenomes of HsAP and HsAqS (Table 9). Without considering normalization, GAL (51.9 Mb) and MUP (14.2 Mb) had the highest and lowest number of SEED assigned reads respectively. Yet, the highest percentage of functionally annotated reads was retrieved from BLA (28%) followed by DAL (26%), respectively and the lowest was from MUP (8.5%).

In this study, MEGAN-SEED scanned the whole collection of contigs to find ORFs assigned to the subsystem connections of carbohydrate, nitrogen, phosphorous and sulfur metabolisms (Table 9, Appendix 3-8). Rhizobiales, Burholderiales and Propionibacteriales, are among the top five abundant orders in HsAPs, which are the primary sources of ORFs for predicted protein families in the carbohydrate, nitrogen, phosphorus, and sulfur metabolism (Appendix 9). In DAL and BLA, reads assigned to the above three taxonomic orders covered more than 70% and 67%, respectively for functional annotation to the four nutrient metabolisms. For GAL, reads assigned in Rhizobiales, Burkholderiales and Propionibacteriales covered approximately 70% of binning to carbohydrate, nitrogen and sulfur metabolisms, while only less than 30% was assigned to phosphorous metabolism.

In the cases of HsAqS, MEGAN assigned more than 4.5 Mb into the top three most abundant orders (Haloferacales, Halobacteriales, identified and/or unidentified orders of Bacteroidetes). Yet, reads from these three orders only covered less than 50% of functional annotation to the four nutrient metabolisms (Fig. 14B and Appendix 9).

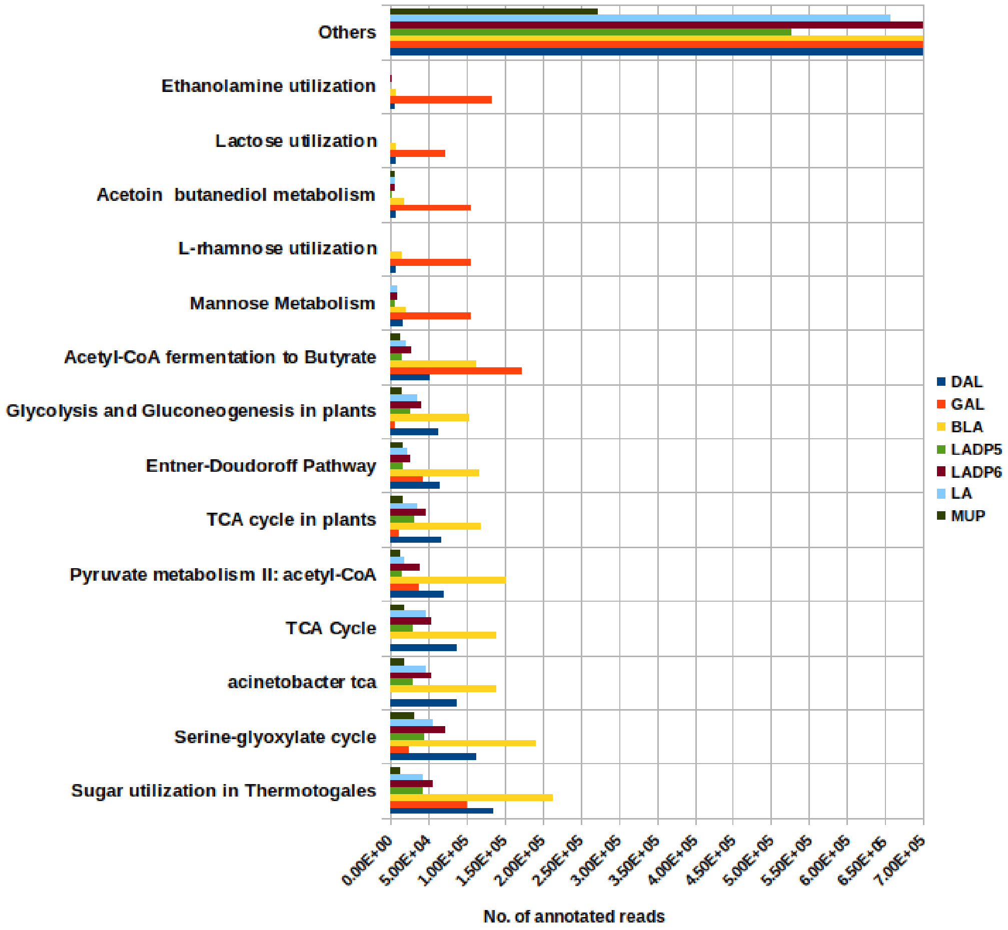
Table 9 Functional annotation of reads based on SEED database

	DAL	**GAL**	**BLA**	**LADP5**	**LADP6**	**LA**	**MUP**
Total assigned reads using SEED (absolute)	27.3 Mb	51.9 Mb	45.4 Mb	17.5 Mb	26.1 Mb	22.6 Mb	14.2 Mb
Total functionally annotated reads (absolute)	7 Mb	9.2 Mb	12.7 Mb	1.9 Mb	3.7 Mb	2.5 Mb	1.2 Mb
Total no. of predicted genes	2200	1100	2600	800	1200	900	700
Total aligned bases assigned in Carbohydrate metabolism	837 Kb	1.35Mb	1.44 Mb	230 Kb	430 Kb	300 Kb	130 Kb
No. of predicted genes in Carbohydrate metabolism	430	230	570	150	200	170	150
Total aligned bases assigned in Nitrogen metabolism	105 Kb	165 Kb	215 Kb	13 Kb	33 Kb	24 Kb	10 Kb
No. of predicted genes in Nitrogen metabolism	68	27	79	16	20	20	12
Total aligned bases assigned in Phosphorus metabolism	76.8 Kb	2.86 Kb	132 Kb	35 Kb	72 Kb	64 Kb	29 Kb
No. of predicted genes in Phosphorus metabolism	40	4	42	16	22	16	33
Total aligned bases assigned in Sulfur metabolism	113 Kb	138 Kb	198 Kb	25 Kb	52 Kb	29 Kb	15 Kb
No. of predicted genes in	59	16	72	17	23	19	14

Generally, the total functionally annotated reads are higher in the HsAP than in HsAqS and predicted protein families in carbohydrate, nitrogen, sulfur and phosphorus metabolisms (Appendix 5-8). High number of protein families involved in carbohydrate metabolism were predicted in all assembled metagenomes studied. From the three HsAPs, the least number of protein coding genes involved in the four studied nutrient metabolisms were predicted in GAL, especially for phosphorous metabolism (Table 9). In the case of HsAqS, relatively similar but small number of genes were predicated for nitrogen, sulfur and phosphorous metabolisms, while more protein-coding genes were predicated for carbohydrate metabolism (Table 9).

Considering only carbohydrate metabolism, approximately 11.0% of annotated reads from DAL and BLA and 15.0% from GAL metagenomes were categorized within the subsystems. For the coastal sites of Lake As’ale metagenomes, average of 11.5% of the annotated reads were grouped under subsystems of carbohydrate metabolism while for MUP the amount was approximately 11.0%. In all the studied sampling sites, 129 different subsystems were involved in predicting carbohydrate metabolism (Appendix 5). The constructed spreadsheet of subsystems for carbohydrate metabolisms indicated variations between the three HsAP and the two HsAqS. For DAL, BLA, MUP and Lake As’ale metagenomes, the majority of these sequences were linked to Sugar utilization in Thermotogales, Serine-glyoxylate cycle, Acinetobacter TCA and TCA Cycle subsystems. For GAL however, Acetyl-CoA fermentation to Butyrate, Mannose Metabolism and Ethanolamine utilization are the three major subsystems involved in carbohydrate metabolism (Fig. 15).

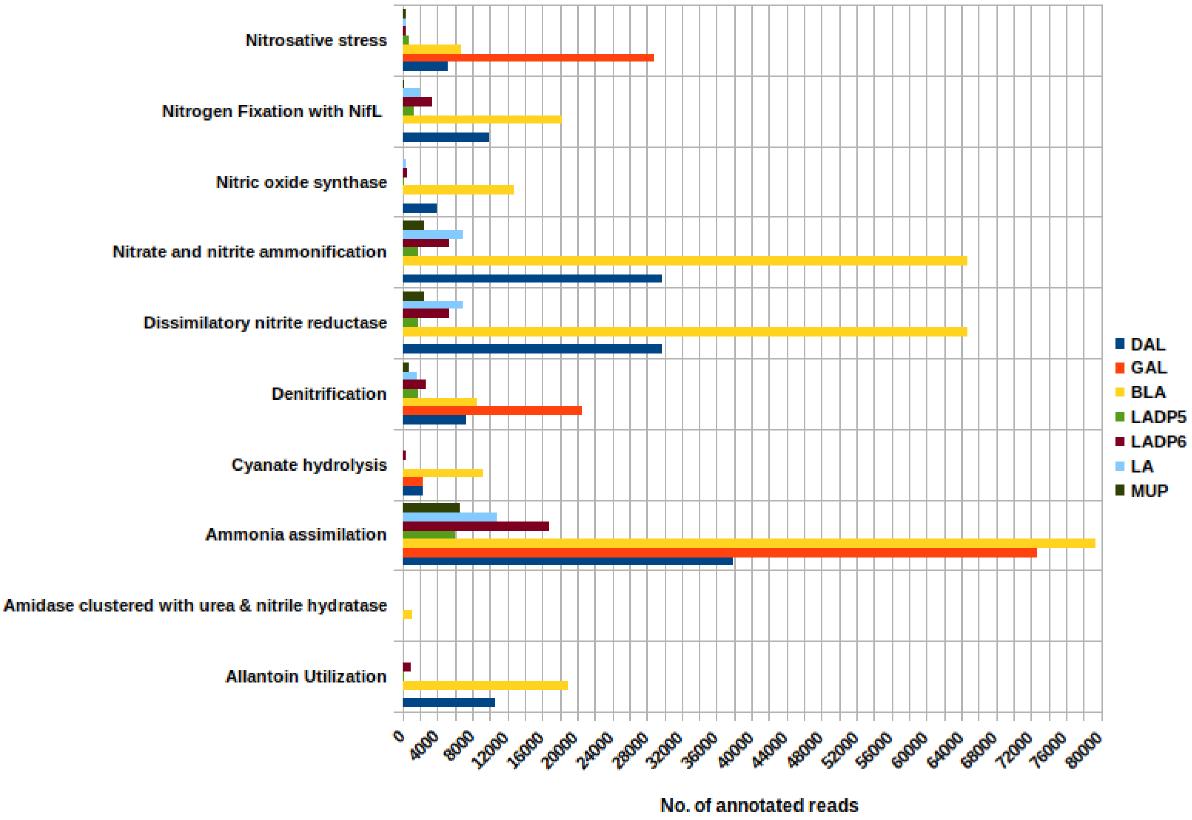
Protein families in the subsystems ketogluconate utilization, lactose utilization, tricarballylate utilization, malonate decarboxylase, D-galactonate catabolism, D-sorbitol (D-Glucitol) and L-sorbose utilization, and inositol catabolism were only predicted in the three HsAP. DAL and BLA exhibited similarities in the types of predicted subsystems such as; D-allose utilization, soluble methane monooxygenase (sMMO), erythritol utilization, sucrose utilization, L-fucose utilization, beta-glucuronide utilization, beta-glucoside metabolism, 2-O-alpha-mannosyl-D-glycerate utilization and others. However, these subsystems were not predicted in GAL as in the case of the other two HsAP. From the three HsAP, GAL showed difference in the types of subsystems as well as predicted protein families. For example, ample translated reads were parsed under the subsystem of methanogenesis from methylated compounds, which is exclusively predicted in GAL (Appendix 5). The Acetone carboxylase was the only subsystem in the carbohydrate metabolism that was exclusively predicted in HsAqS.



Major SEED subsystem-connections for carbohydrate metabolism in metagenomes of DAL (blue), GAL (orange), BLA (yellow), Lake As'ale coastal sites (green, brown and light blue) and MUP (black) datasets determined using DIAMOND-MEGAN analysis.

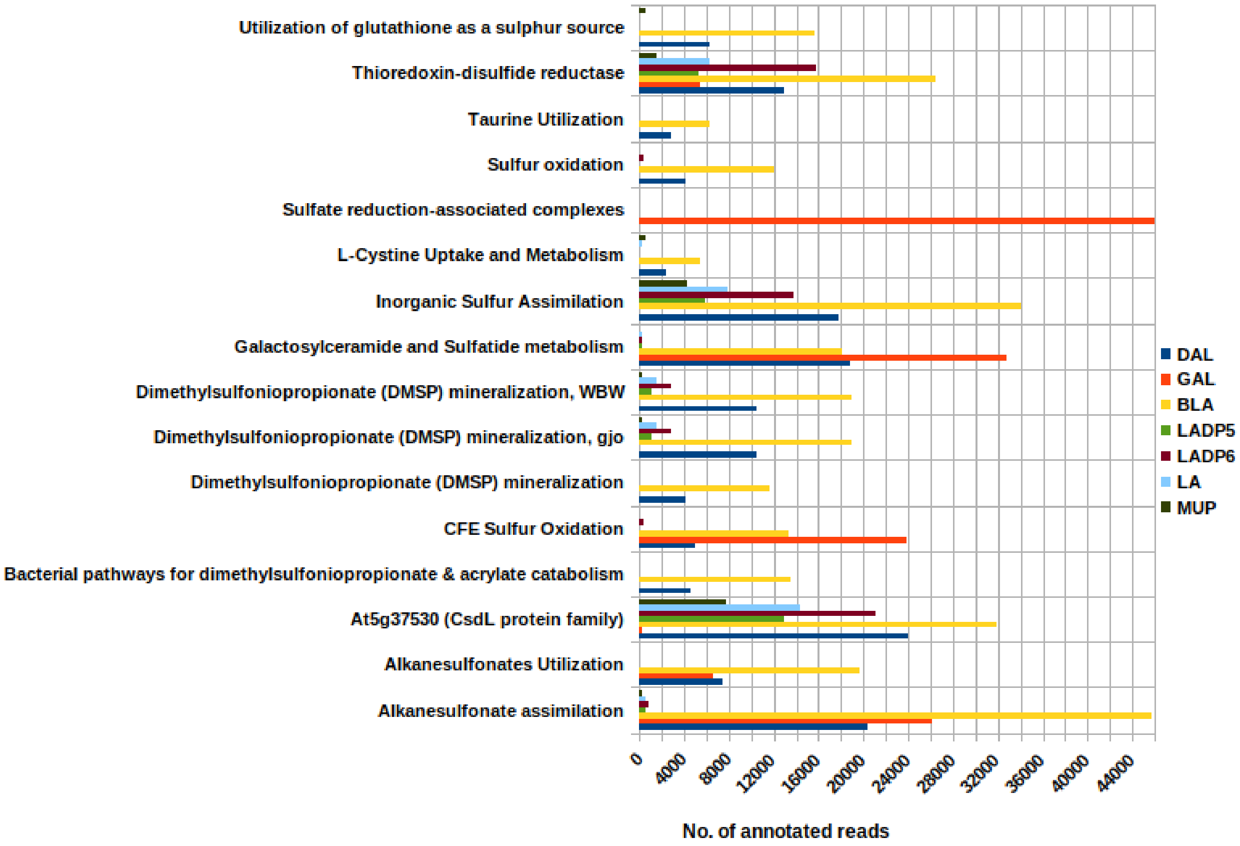
In addition to the different subsystems of carbohydrate metabolism, total of 10 subsystems in nitrogen metabolism, 6 subsystems in phosphorous metabolism and 16 subsystems in sulfur metabolism were predicted in the five studied sites (Fig 16-18). In this investigation, higher number of subsystem connections were predicated to nitrogen and sulfur metabolisms than to phosphorous metabolism. For nitrogen and sulfur metabolisms, an approximate range of 1.3% to 1.8% and 1.1% to 1.6% of annotated reads were categorized in the predicted subsystems for the three HsAP and MUP. For the sample sites of Lake As’ale, less than 1% of annotated reads were grouped in subsystems of nitrogen metabolism, while an average of 1.3% of annotated reads was categorized for sulfur metabolism (Table 9).

In metagenomes of DAL and BLA, more than 90% of subsystems parsed by SEED belonged to the predicted metabolisms of nitrogen, sulfur and phosphorous (Fig 16-18). However, for GAL, the number of protein coding genes predicted for nitrogen metabolism was minimum (2.45% of the total) and only 4 out of the 10 subsystems were predicted (Fig. 16). The highest amount of total annotated reads and gene coding ORFs categorized to subsystems of phosphate metabolism were obtained from MUP metagenome (2.4% and 4.9%, respectively) and the lowest was obtained from GAL metagenome (0.03% and 0.4%, respectively) (Table 9). In general, the total amount of predicated protein families was reflected by the the percentage of functionally annotated reads.

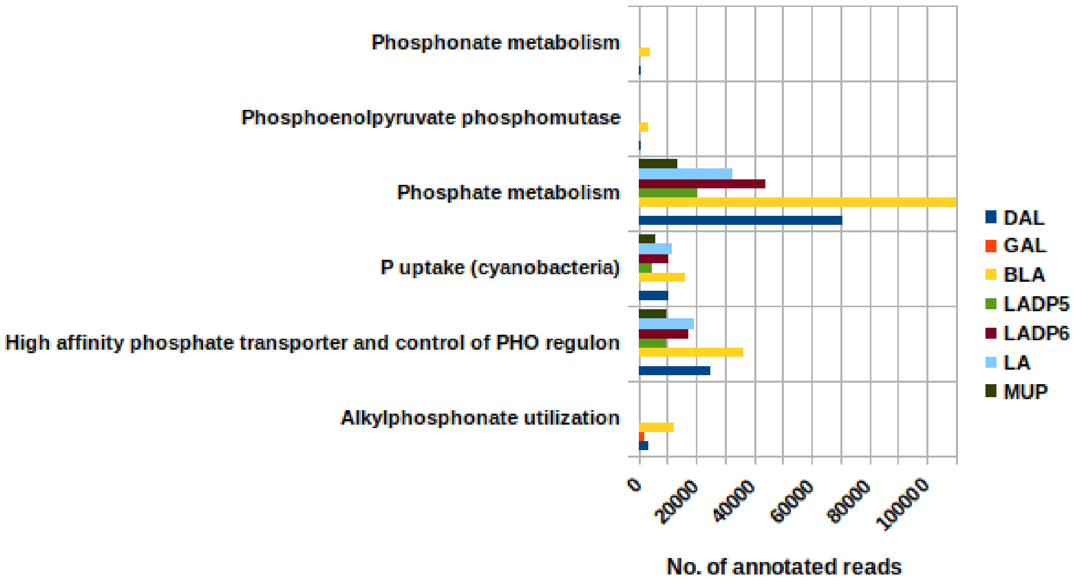


Major SEED subsystem-connections for nitrogen metabolism in metagenomes of DAL (blue) GAL (orange), BLA (yellow), Lake As'ale coastal sites (green, brown and light blue) and MUP (black) datasets determined using DIAMOND-MEGAN analysis.

Furthermore, according to SEED annotation, fewer and sometimes rare protein families involved in nitrogen, sulfur and phosphorous metabolisms were predicted for HsAqS and GAL (Appendixes 6-8). For example, protein families in GAL, such as those belonging to the subsystem 'sulfate reduction-associated complexes', were neither observed in DAL and BLA nor in the HsAqS. Then again, many protein families the subsystems such as 'dimethylsulfoniopropionate (DMSP) mineralization', 'L-cystine uptake and metabolism', 'taurine utilization', ‘phosphoenolpyruvate phosphomutase’, ‘amidase clustered with urea and nitrile hydratase functions’ and so on were mostly if not exclusively predicted in DAL and BLA than the rest of the studied sampling sites (Appendixes 6-8). However, the number of protein families predicted for phosphorous cycle was minimal and couldn’t generate clear map of its metabolic pathway in all sampling sites.



Major SEED subsystem-connections for sulfur metabolism in of DAL (blue), GAL (orange), BLA (yellow), Lake As’ale coastal sites (green, brown and light blue) and MUP (black) datasets determined using DIAMOND-MEGAN analysis.



Major SEED subsystem connections for phosphate metabolism in metagenomes of DAL (blue), GAL (orange), BLA (yellow), Lake As'ale coastal sites (green, brown and light blue) and MUP (black) datasets determined using DIAMOND-MEGAN analysis.

4.3.3. Mapping of Carbon fixation in HsAqS and HsAP

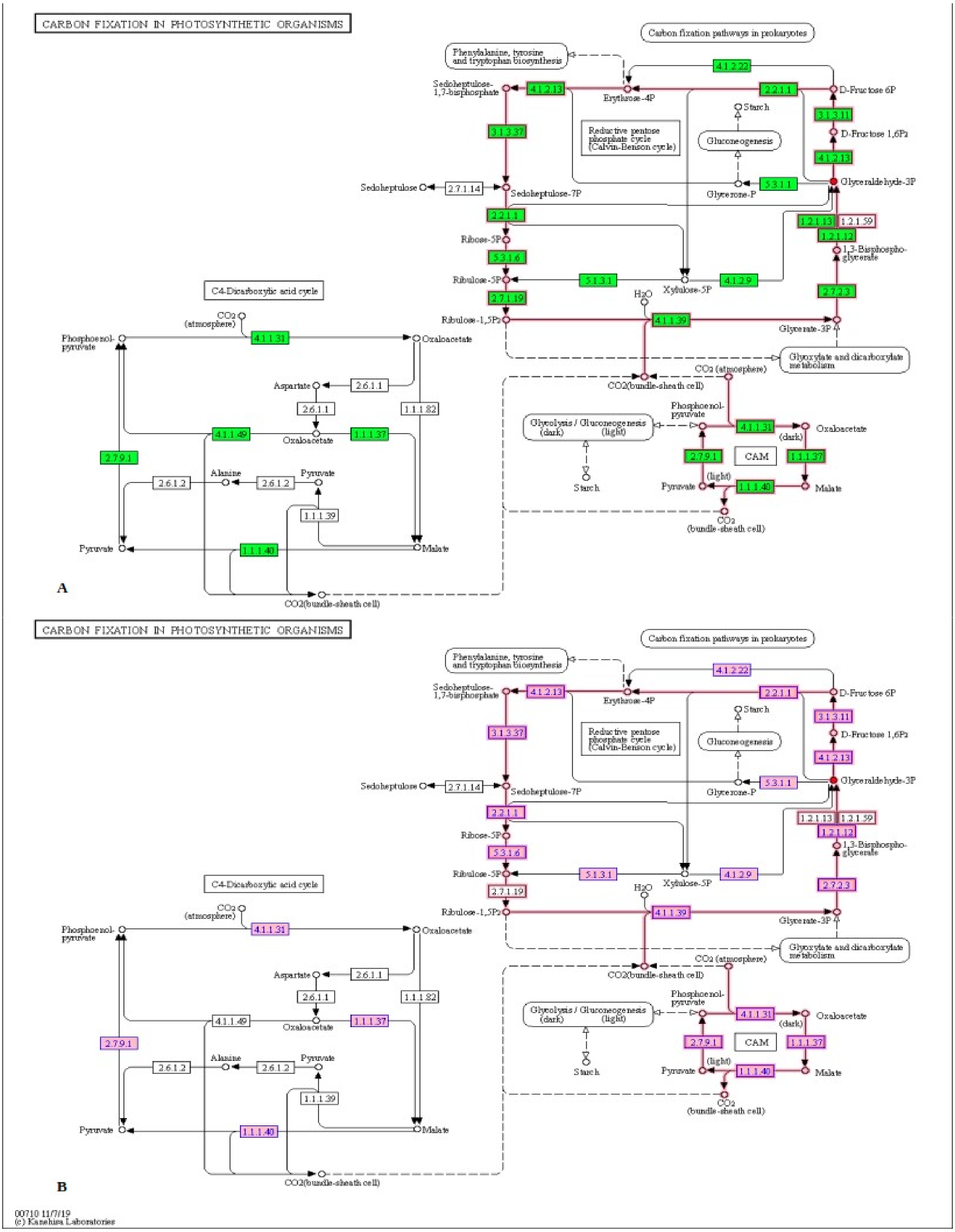
The SEED annotated twelve different subsystems involved in carbon fixation in all sampling sites (Table 10). Furthermore, KEGG Orthology (KO) numbers were assigned for the 100 predicted protein coding genes (total predicated genes in carbon fixation was 107). The KO numbers were eventually used for mapping specific metabolic pathways with KEGG (Table 10). For DAL and BLA large number of proteins with their associated KO (84 and 83, respectively) were predicted, while the least was predicted in GAL (i.e., 26). In HsAqS metagenomes, the protein NAD(P)-

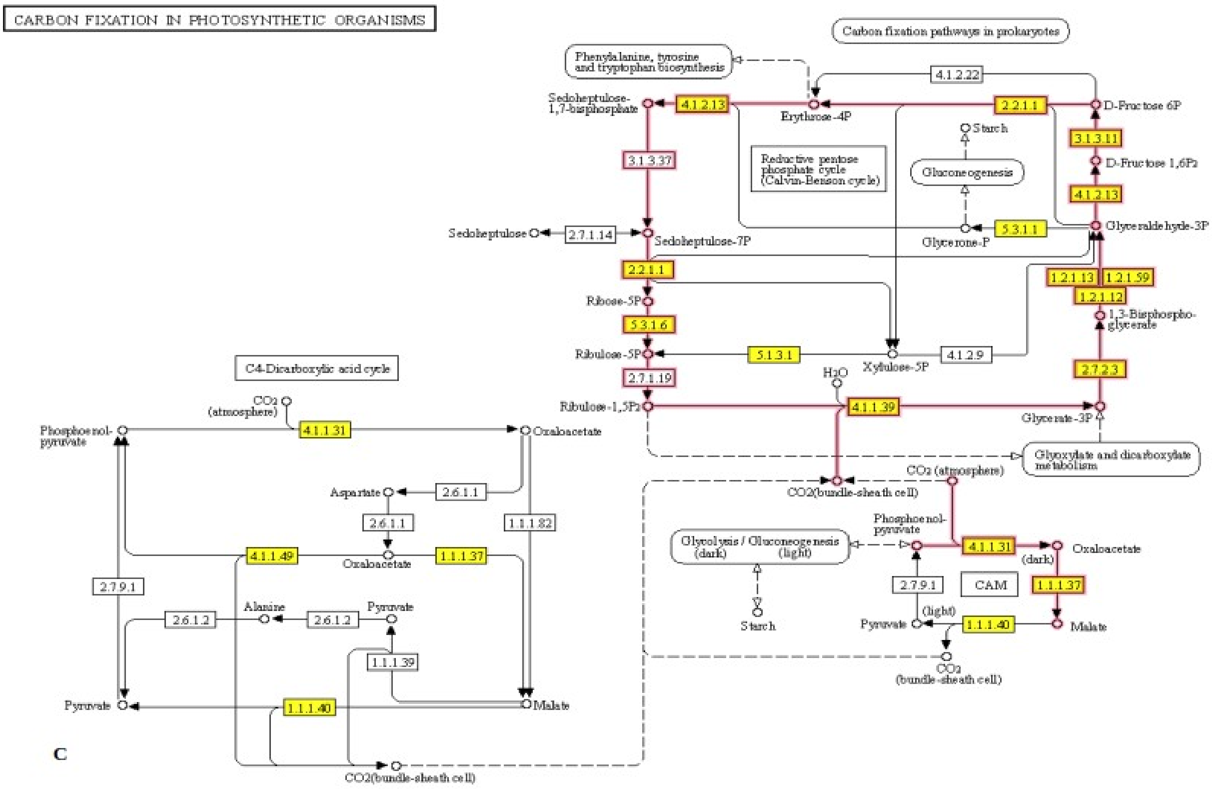
dependent glyceraldehyde 3-phosphate dehydrogenase archaeal (EC:1.2.1.59) and Isocitrate dehydrogenase [NAD] (EC 1.1.1.41) were uniquely predicted. Proteins such as malate synthase G (EC:2.3.3.9), serine--glyoxylate aminotransferase (EC:2.6.1.45), ferredoxin-dependent glutamate synthase (EC:1.4.7.1), glycolate oxidase (EC:1.1.3.15) were only predicated in the three HsAP. However proteins such as fumarate hydratase class I, aerobic (EC:4.2.1.2), fructose-bisphosphate aldolase class II (EC:4.1.2.13), glyoxylate carboligase (EC:4.1.1.47), aconitate hydratase 2 (EC:4.2.1.3), chloride channel protein, putative dihydrolipoamide dehydrogenase (EC:1.8.1.4), glycolate dehydrogenase (EC:1.1.99.14), FAD-binding subunit GlcE, malate synthase (EC:2.3.3.9), transketolase, C-terminal section (EC:2.2.1.1), malyl-CoA lyase (EC:4.1.3.24) and so on were exclusively predicated in DAL and BLA metagenomes (Appendix 5). Therefore, based on SEED predication of genes, complete cycle of reductive pentose phosphate cycle and crassulacean acid metabolism (dark or light reactions) were mapped for DAL, BLA (Fig. 19 A and B, respectively) and for coastal sites of Lake As’ale and MUP (Fig. 19 C-F). However, for GAL, no complete metabolic paths involved in carbon fixation was mapped.

Table 10 SEED functional annotation of aligned bases to subsystems in carbon fixation in kilo bases (kb)

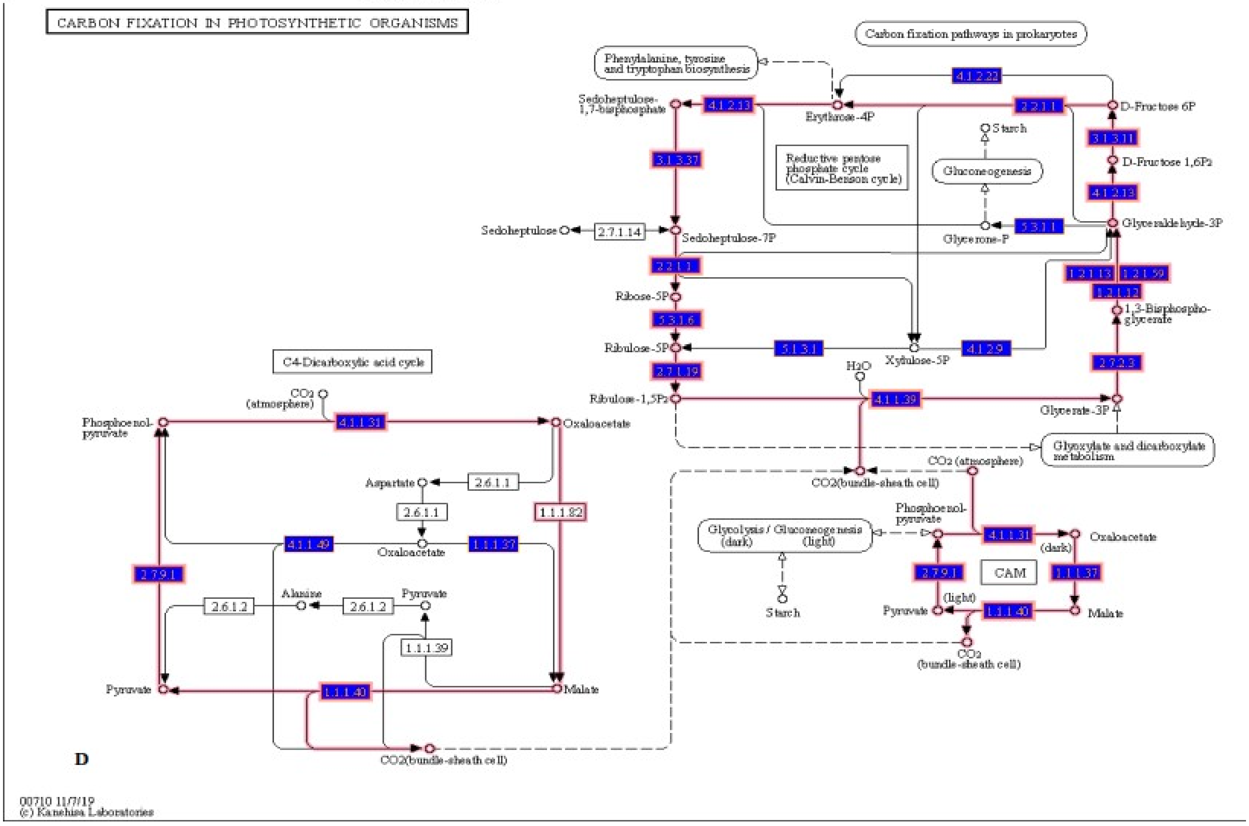
Subsystems in Carbon fixation	**HsAP**			**HsAqS**			
	DAL	**GAL**	**BLA**	**LADP5**	**LADP6**	**LA**	**MUP**
Calvin-Benson cycle	26.3	26	40.4	9.5	15.5	11.8	4.5

Calvin-Benson-Bassham cycle	15	25.8	23.9	4.4	8.3	7.2	3.1
Carboxysome	5.7	34.8	12.7	1.7	2.9	4.5	0.4
CO_2 uptake carboxysome	4.7	18.1	10.8	3.2	7.7	8.2	1.9
Ethylmalonyl-CoA pathway of C_2 assim.	11.1	18.5	18.9	0.3	2.8	0.9	2.5
Ethylmalonyl-CoA pathway of C_2 assim., GJO	11.1	18.5	19	0.3	2.8	0.9	2.5
Pentose phosphate pathway	24.2	41.3	40.6	11.5	17.5	12.3	3.1
Pentose phosphate pathway in plants	17.8	40.6	25.1	7.2	13.2	8.1	2.7
Photorespiration (oxidative C_2 cycle)	47.6	15.6	84.6	13.4	22.4	18.9	10.9
Photorespiration (oxidative C_2 cycle)plants	27.3	32	44.9	12.2	16.4	14.7	6.7
TCA cycle	86.3	0.8	139.5	30.7	53.5	46.8	19.3
TCA cycle in plants	65.2	11.7	118.7	32.3	46.4	35.6	16.5
Total No. of protein coding genes	85	26	86	46	58	54	39
Total No. of KO	84	26	83	45	56	50	38


CARBON FIXATION IN PHOTOSYNTHETIC ORGANISMS
Carbon fixation pathways in prokaryotes
Phenylalanine, tyrosine and tryptophan biosynthesis
Sedoheptulose-1,7-bisphosphate
Erythrose-4P
D-Fructose 6P
D-Fructose 1,6P2
Glyceraldehyde-3P
Starch
Reductive pentose phosphate cycle (Calvin-Benson cycle)
Gluconeogenesis
Glycerone-P
Sedoheptulose
Sedoheptulose-7P
Ribose-5P
Ribulose-5P
Xylulose-5P
Ribulose-1,5P2
1,3-Bisphospho-glycerate
Glycerate-3P
H2O
CO2 (atmosphere)
CO2(bundle-sheath cell)
Glyoxylate and dicarboxylate metabolism
C4-Dicarboxylic acid cycle
Phosphoenol-pyruvate
Oxaloacetate
Aspartate
Alanine
Pyruvate
Malate
Glycolysis / Gluconeogenesis (dark) (light)
CAM
(dark)
(light)
CO2 (bundle-sheath cell)
4.1.2.22
4.1.2.13
2.2.1.1
3.1.3.11
3.1.3.37
5.3.1.1
2.7.1.14
1.2.1.13
1.2.1.59
1.2.1.12
5.3.1.6
5.1.3.1
4.1.2.9
2.7.1.19
2.7.2.3
4.1.1.39
4.1.1.31
2.6.1.1
1.1.1.82
4.1.1.49
1.1.1.37
2.7.9.1
2.6.1.2
1.1.1.39
1.1.1.40
A
B
00710 11/7/19
(c) Kanehisa Laboratories


CARBON FIXATION IN PHOTOSYNTHETIC ORGANISMS
Carbon fixation pathways in prokaryotes
Phenylalanine, tyrosine and tryptophan biosynthesis
Sedoheptulose-1,7-bisphosphate
4.1.2.13
Erythrose-4P
4.1.2.22
2.2.1.1
D-Fructose 6P
3.1.3.11
D-Fructose 1,6P2
4.1.2.13
Glyceraldehyde-3P
Starch
Gluconeogenesis
Reductive pentose phosphate cycle (Calvin-Benson cycle)
3.1.3.37
Sedoheptulose
2.7.1.14
Sedoheptulose-7P
Glycerone-P
5.3.1.1
2.2.1.1
1.2.1.13
1.2.1.59
1.2.1.12
Ribose-5P
1,3-Bisphosphoglycerate
5.3.1.6
Ribulose-5P
5.1.3.1
Xylulose-5P
4.1.2.9
2.7.1.19
H2O
2.7.2.3
Ribulose-1,5P2
4.1.1.39
Glycerate-3P
Glyoxylate and dicarboxylate metabolism
C4-Dicarboxylic acid cycle
CO2 (atmosphere)
Phosphoenol-pyruvate
4.1.1.31
Oxaloacetate
CO2(bundle-sheath cell)
CO2 (atmosphere)
Aspartate
2.6.1.1
2.6.1.1
1.1.1.82
Phosphoenol-pyruvate
Glycolysis / Gluconeogenesis (dark) (light)
4.1.1.31
Oxaloacetate
4.1.1.49
Oxaloacetate
1.1.1.37
(dark)
2.7.9.1
CAM
1.1.1.37
2.7.9.1
Alanine
Pyruvate
2.6.1.2
2.6.1.2
Starch
(light)
Pyruvate
1.1.1.40
Malate
1.1.1.39
CO2 (bundle-sheath cell)
Pyruvate
1.1.1.40
Malate
CO2(bundle-sheath cell)

C


CARBON FIXATION IN PHOTOSYNTHETIC ORGANISMS
Carbon fixation pathways in prokaryotes
Phenylalanine, tyrosine and tryptophan biosynthesis
4.1.2.22
Sedoheptulose-1,7-bisphosphate
4.1.2.13
Erythrose-4P
2.2.1.1
D-Fructose 6P
Starch
3.1.3.11
Reductive pentose phosphate cycle (Calvin-Benson cycle)
Gluconeogenesis
3.1.3.37
D-Fructose 1,6P2
4.1.2.13
Sedoheptulose
2.7.1.14
Sedoheptulose-7P
5.3.1.1
Glyceraldehyde-3P
Glycerone-P
2.2.1.1
1.2.1.13
1.2.1.59
1.2.1.12
Ribose-5P
1,3-Bisphosphoglycerate
5.3.1.6
Ribulose-5P
5.1.3.1
4.1.2.9
Xylulose-5P
C4-Dicarboxylic acid cycle
2.7.1.19
H2O
2.7.2.3
CO2 (atmosphere)
Ribulose-1,5P2
4.1.1.39
Glycerate-3P
Phosphoenol-pyruvate
4.1.1.31
Oxaloacetate
Glyoxylate and dicarboxylate metabolism
CO2 (atmosphere)
Aspartate
2.6.1.1
CO2(bundle-sheath cell)
2.6.1.1
1.1.1.82
Phosphoenol-pyruvate
Glycolysis / Gluconeogenesis (dark) (light)
4.1.1.31
Oxaloacetate
4.1.1.49
Oxaloacetate
1.1.1.37
(dark)
2.7.9.1
2.7.9.1
CAM
1.1.1.37
Alanine
Pyruvate
2.6.1.2
2.6.1.2
Starch
(light)
Pyruvate
1.1.1.40
Malate
1.1.1.39
CO2 (bundle-sheath cell)
Pyruvate
1.1.1.40
Malate
CO2(bundle-sheath cell)
00710 11/7/19
(c) Kanehisa Laboratories

D

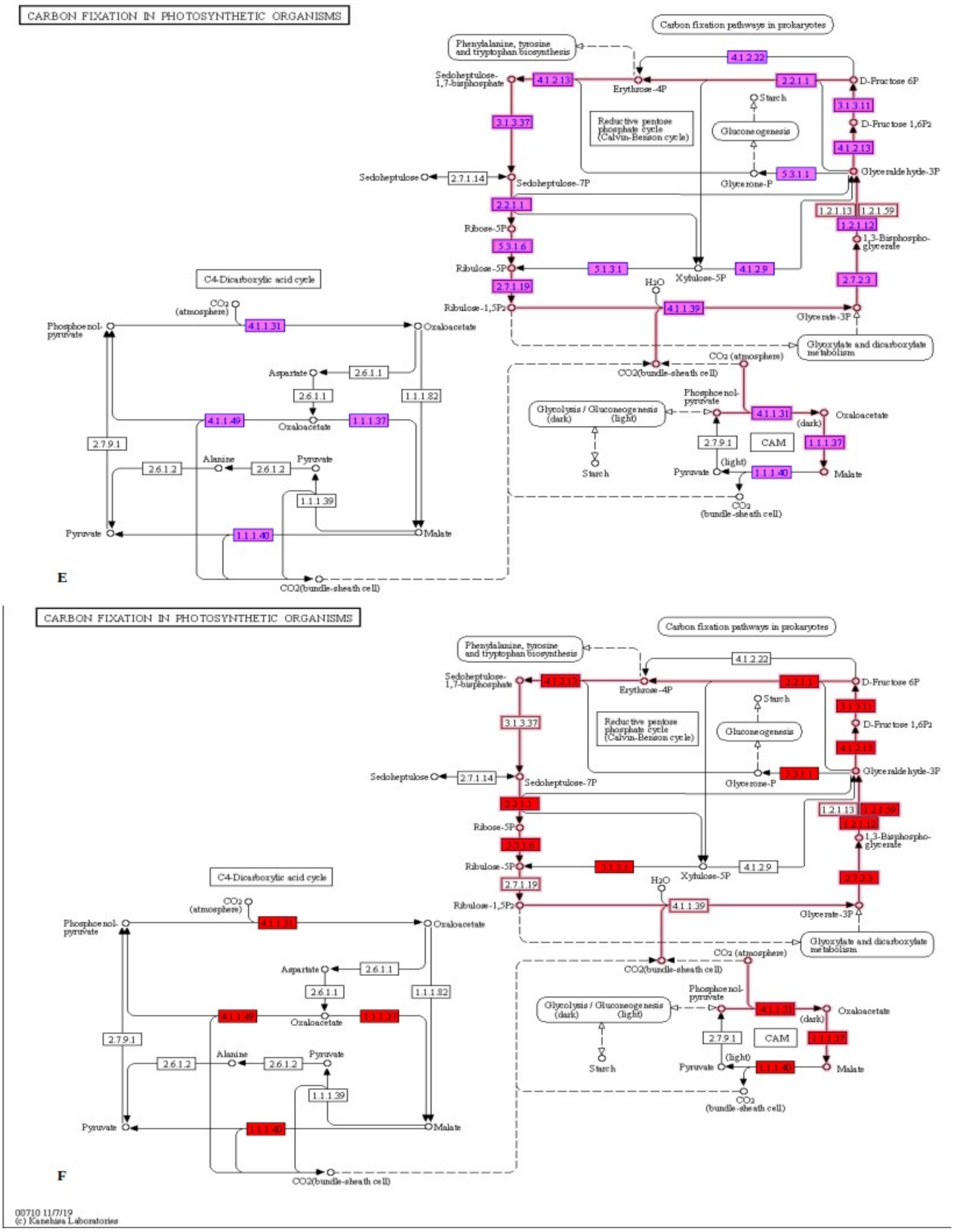


Fig. 19 Schematic of metabolic reconstruction of carbon fixation

Carbon fixation based on the predicted ORFs using SEED- KEGG. DAL [A], BLA [B], LADP5 [C], LADP6 [D], LA [E] and MUP [F].

4.3.4. Mapping of methane metabolism in GAL

Proteins involved in subsystems of “methanogenesis” and “methanogenesis from methylated compounds” were mainly predicted in GAL metagenome (Fig. 20). All of the predicted genes for these subsystems were associated with bacteria, especially with the taxonomic orders Rhizobiales and Burkholderiales (Appendix 11). Only F420-dependent N(5), N(10)-methylenetetrahydromethanopterin reductase (EC:1.5.99.11) and N(5),N(10)-methenyltetrahydromethanopterin cyclohydrolase (EC:3.5.4.27) were absent from GAL’s list of proteins predicated. Hence, complete and/or near complete methane metabolism with four possible methanogenesis pathways were mapped in GAL using KEGG.

The first predicted pathway of methanogenesis starts from CO_2 being reduced to Formyl-MFR by Formylmethanofuran dehydrogenase (EC:1.2.7.12) and then subsequently to methane through many cascades of catalytic reactions by many enezymes (Fig. 21A). Two enzymes that were not detected in GAL metagenome but vital for this pathway are methenyltetrahydromethanopterin cyclohydrolase (EC:3.5.4.27) and 5,10-methylenetetrahydromethanopterin reductase (EC:1.5.98.2). For the second predicted pathway of methanogenesis, two important enzymes acetyl-CoA synthetase (EC:6.2.1.1) and acetyl-CoA decarbonylase/synthase, CODH/ACS complex subunit beta (EC:2.3.1.169) are missing (Fig 21B). Acetyl-CoA synthetase

(EC:6.2.1.1) is important for direct conversion of acetate to acetyl-coA while acetyl-CoA decarbonylase/synthase is important to convert acetyl-CoA to 5-methyltetrahydromethanopterin. Acetyl-CoA synthetase (EC:6.2.1.1) is mainly important in carbon fixation than methanogenesis as it is the final converter of acetate to acetyl coA to enter in all major carbon fixing pathways (reductive citrate cycle, 3-hydroxypropionate bi-cycle, hydroxypropionate-hydroxybutylate cycle, dicarboxylate-hydroxybutyrate cycle, reductive acetyl-CoA pathway, phosphate acetyltransferase-acetate kinase pathway).

In the third predicted pathway of methanogenesis, the main enzymes involved in conversion of methane to methanol (methane/ammonia monooxygenase [EC:1.14.13.25]) or from methanol to 2-(methylthio) ethanesulfonate (methyl-Co (III) methanol-specific corrinoid protein]: coenzyme M methyltransferase [EC:2.1.1.246]) were detected (Fig. 21C). For the fourth predicated pathway of methanogenesis, enzymes involved in the dimethylamine and methylamine metabolisms that produces methyl CoM, that is an intermediate product, were detected (Fig. 21D).

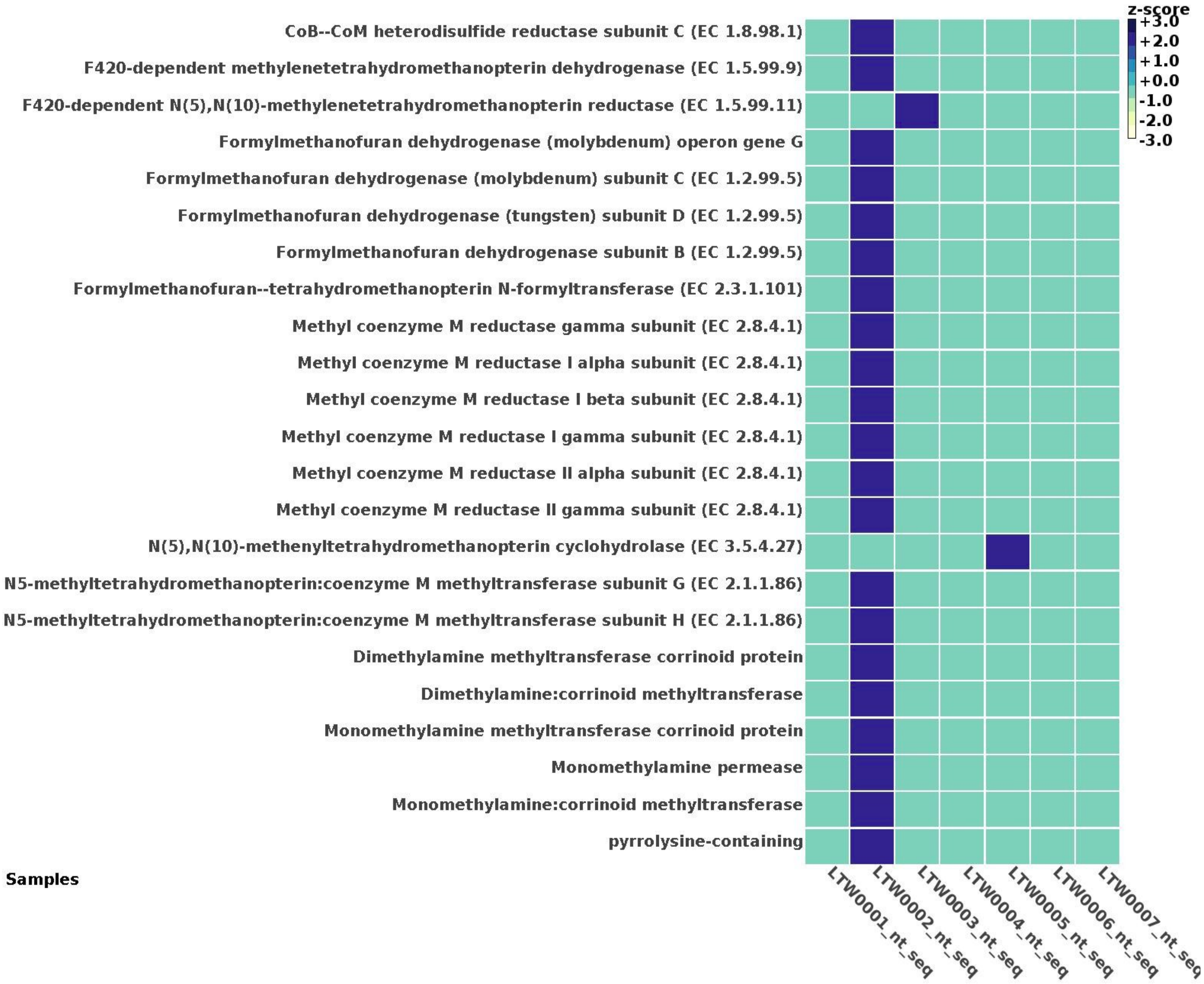
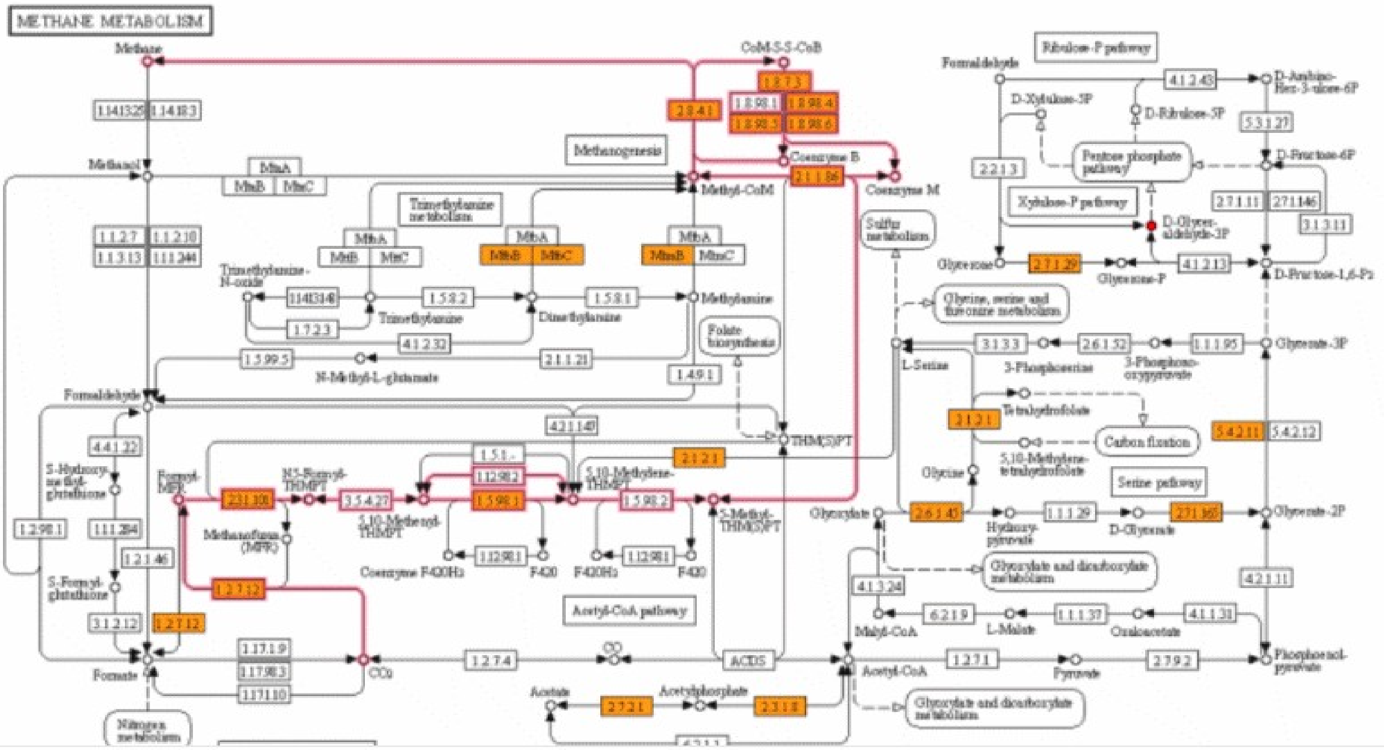
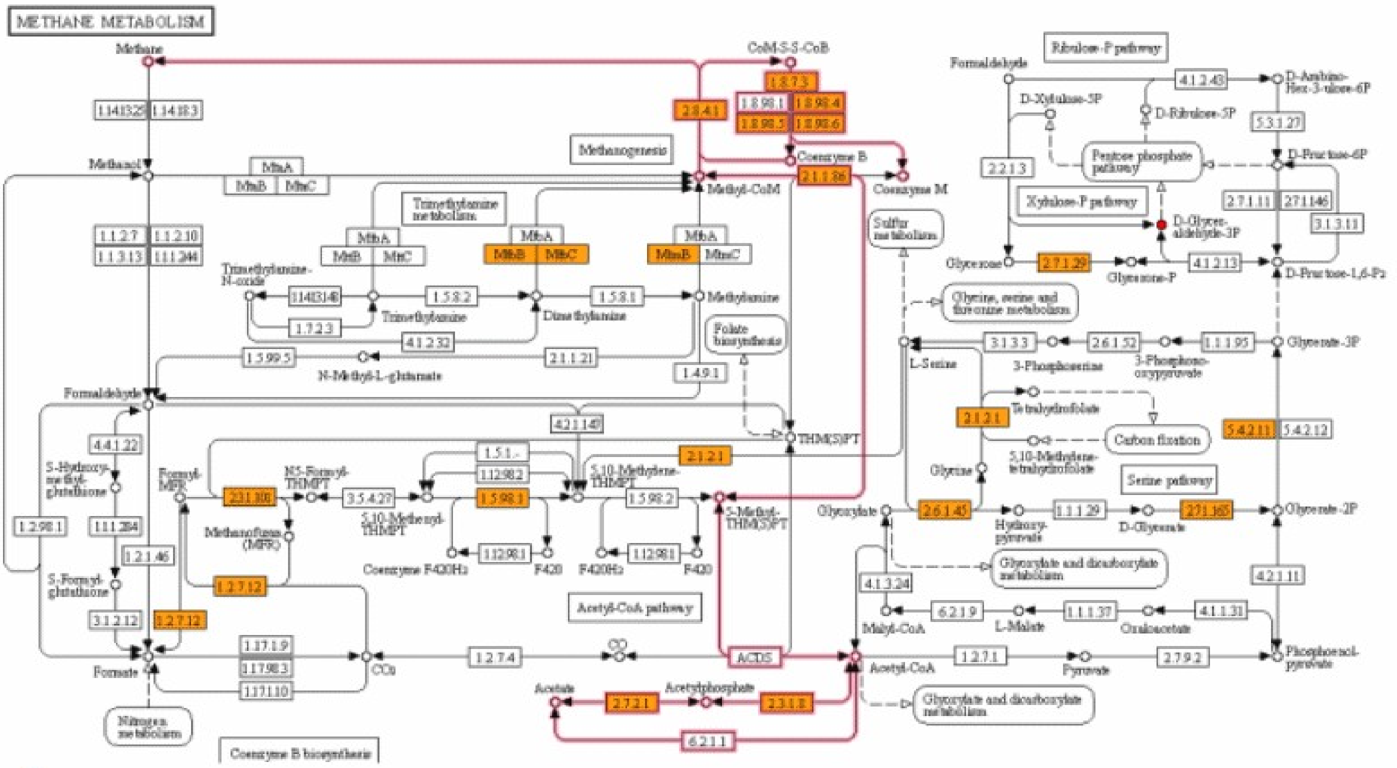


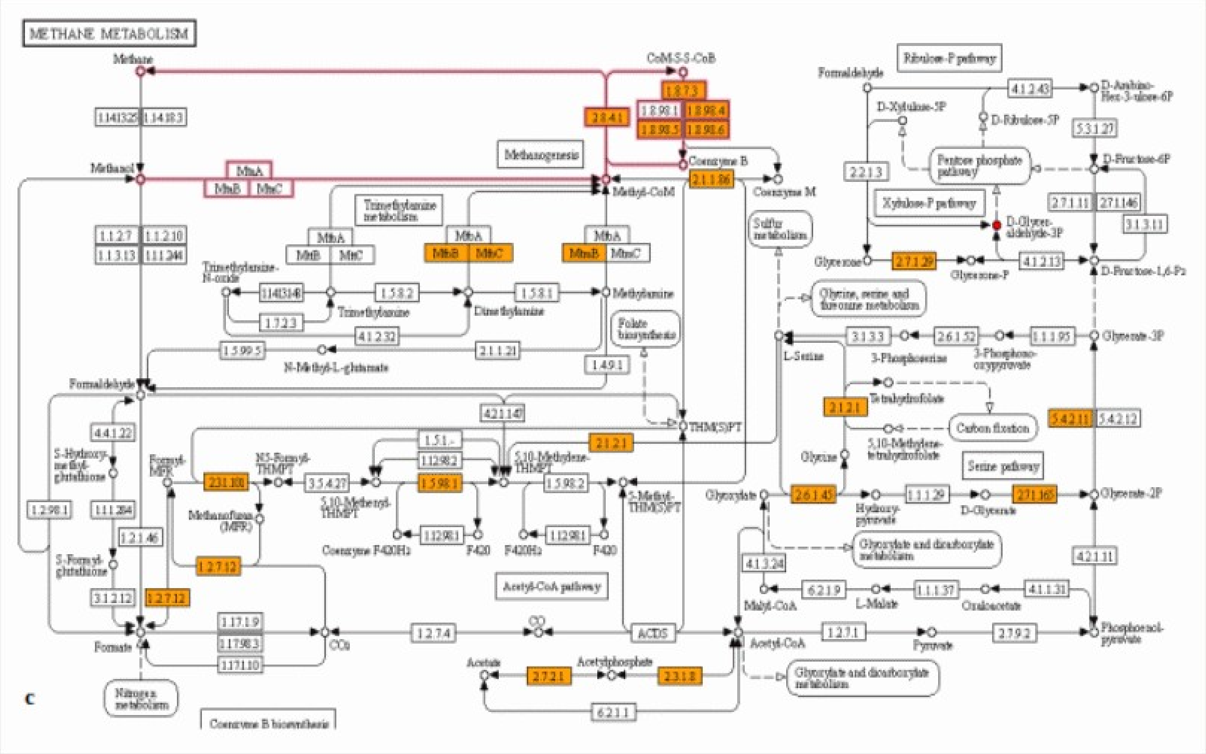
Fig. 20. Heatmap of predicted proteins involved in methanogenesis


METHANE METABOLISM
Methane
CoM-S-S-CoB
1.8.7.3
2.8.4.1
1.8.98.1
1.8.98.4
1.8.98.5
1.8.98.6
1.14.13.25
1.14.18.3
Methanol
MtaA
MtaB
MtaC
Methanogenesis
Coenzyme B
2.1.1.86
Methyl-CoM
Coenzyme M
Trimethylamine metabolism
MttB
MttC
MtbA
MtbB
MtbC
MtmB
MtmC
1.1.2.7
1.1.2.10
1.1.3.13
1.1.1.244
Trimethylamine-N-oxide
1.14.13.148
1.5.8.2
1.5.8.1
Trimethylamine
Dimethylamine
Methylamine
1.7.2.3
4.1.2.32
1.5.99.5
2.1.1.21
N-Methyl-L-glutamate
Folate biosynthesis
1.4.9.1
Formaldehyde
4.2.1.147
THM(S)PT
4.4.1.22
S-Hydroxy-methyl-glutathione
Formyl-MFR
2.3.1.101
N5-Formyl-THMPT
3.5.4.27
1.5.1.-
1.12.98.2
1.5.98.1
5,10-Methenyl-THMPT
5,10-Methylene-THMPT
1.5.98.2
2.1.2.1
5-Methyl-THM(S)PT
1.2.98.1
1.1.1.284
1.2.1.46
Methanofuran (MFR)
1.2.7.12
1.12.98.1
Coenzyme F420H2
F420
F420H2
S-Formyl-glutathione
3.1.2.12
Acetyl-CoA pathway
1.17.1.9
1.17.98.3
1.17.1.10
Formate
CO2
1.2.7.4
CO
ACDS
Acetyl-CoA
Nitrogen metabolism
Acetate
2.7.2.1
Acetylphosphate
2.3.1.8
6.2.1.1
Ribulose-P pathway
Formaldehyde
4.1.2.43
D-Arabino-Hex-3-ulose-6P
D-Xylulose-5P
D-Ribulose-5P
5.3.1.27
2.2.1.3
Pentose phosphate pathway
D-Fructose-6P
Xylulose-P pathway
2.7.1.11
2.7.1.146
3.1.3.11
D-Glyceraldehyde-3P
Sulfur metabolism
Glycerone
2.7.1.29
Glycerone-P
4.1.2.13
D-Fructose-1,6-P2
Glycine, serine and threonine metabolism
L-Serine
3.1.3.3
2.6.1.52
1.1.1.95
Glycerate-3P
3-Phosphoserine
3-Phosphono-oxypyruvate
2.1.2.1
Tetrahydrofolate
Carbon fixation
5.4.2.11
5.4.2.12
Glycine
5,10-Methylene-tetrahydrofolate
Serine pathway
Glyoxylate
2.6.1.45
1.1.1.29
2.7.1.165
Glycerate-2P
Hydroxy-pyruvate
D-Glycerate
Glyoxylate and dicarboxylate metabolism
4.1.3.24
4.2.1.11
Malyl-CoA
6.2.1.9
L-Malate
1.1.1.37
Oxaloacetate
4.1.1.31
1.2.7.1
Pyruvate
2.7.9.2
Phosphoenol-pyruvate
Glyoxylate and dicarboxylate metabolism

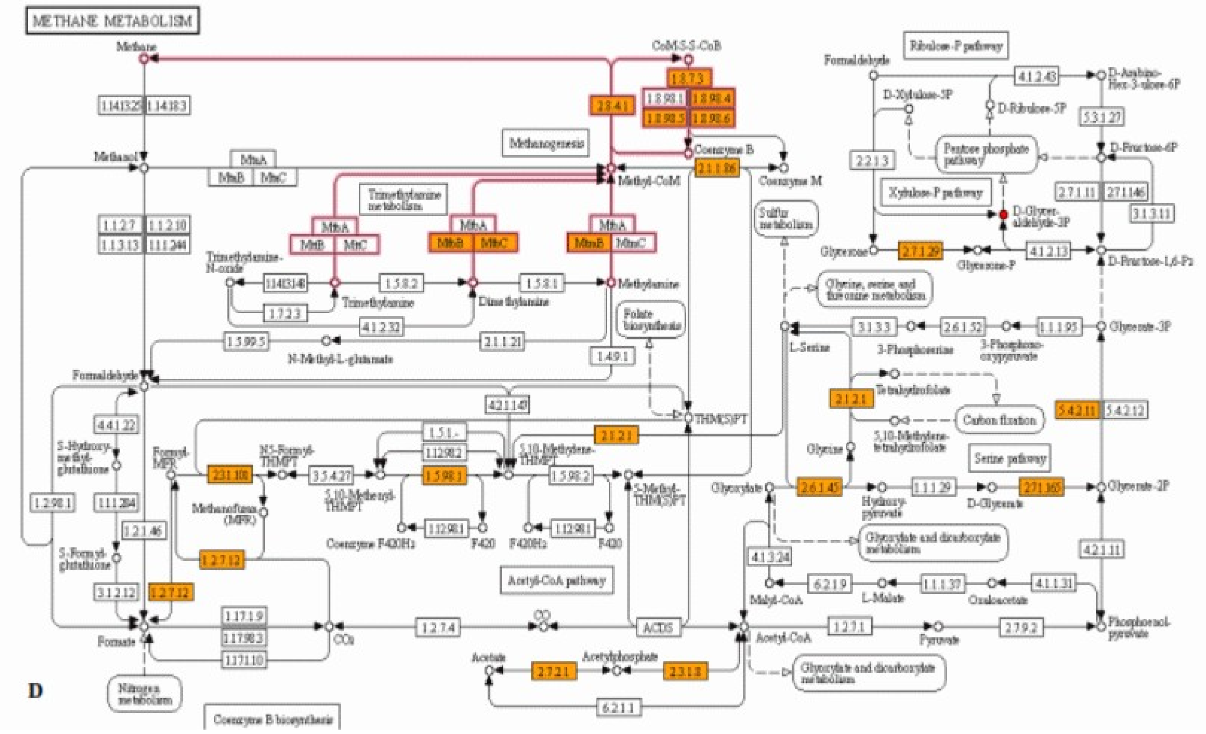
A


METHANE METABOLISM
Methane
CoM-S-S-CoB
1.8.7.3
2.8.4.1
1.8.98.1
1.8.98.4
1.8.98.5
1.8.98.6
1.14.13.25
1.14.18.3
Methanol
MtaA
MtaB
MtaC
Methanogenesis
Coenzyme B
2.1.1.86
Methyl-CoM
Coenzyme M
Trimethylamine metabolism
MttB
MttC
MtbA
MtbB
MtbC
MtmB
MtmC
1.1.2.7
1.1.2.10
1.1.3.13
1.1.1.244
Trimethylamine-N-oxide
1.14.13.148
1.5.8.2
1.5.8.1
Trimethylamine
Dimethylamine
Methylamine
1.7.2.3
4.1.2.32
1.5.99.5
2.1.1.21
N-Methyl-L-glutamate
Folate biosynthesis
1.4.9.1
Formaldehyde
4.2.1.147
THM(S)PT
4.4.1.22
S-Hydroxy-methyl-glutathione
Formyl-MFR
2.3.1.101
N5-Formyl-THMPT
3.5.4.27
1.5.1.-
1.12.98.2
1.5.98.1
5,10-Methenyl-THMPT
5,10-Methylene-THMPT
1.5.98.2
2.1.2.1
5-Methyl-THM(S)PT
1.2.98.1
1.1.1.284
1.2.1.46
Methanofuran (MFR)
1.2.7.12
1.12.98.1
Coenzyme F420H2
F420
F420H2
S-Formyl-glutathione
3.1.2.12
Acetyl-CoA pathway
1.17.1.9
1.17.98.3
1.17.1.10
Formate
CO2
1.2.7.4
CO
ACDS
Acetyl-CoA
Nitrogen metabolism
Coenzyme B biosynthesis
Acetate
2.7.2.1
Acetylphosphate
2.3.1.8
6.2.1.1
Ribulose-P pathway
Formaldehyde
4.1.2.43
D-Arabino-Hex-3-ulose-6P
D-Xylulose-5P
D-Ribulose-5P
5.3.1.27
2.2.1.3
Pentose phosphate pathway
D-Fructose-6P
Xylulose-P pathway
2.7.1.11
2.7.1.146
3.1.3.11
D-Glyceraldehyde-3P
Sulfur metabolism
Glycerone
2.7.1.29
Glycerone-P
4.1.2.13
D-Fructose-1,6-P2
Glycine, serine and threonine metabolism
L-Serine
3.1.3.3
2.6.1.52
1.1.1.95
Glycerate-3P
3-Phosphoserine
3-Phosphono-oxypyruvate
2.1.2.1
Tetrahydrofolate
Carbon fixation
5.4.2.11
5.4.2.12
Glycine
5,10-Methylene-tetrahydrofolate
Serine pathway
Glyoxylate
2.6.1.45
1.1.1.29
2.7.1.165
Glycerate-2P
Hydroxy-pyruvate
D-Glycerate
Glyoxylate and dicarboxylate metabolism
4.1.3.24
4.2.1.11
Malyl-CoA
6.2.1.9
L-Malate
1.1.1.37
Oxaloacetate
4.1.1.31
1.2.7.1
Pyruvate
2.7.9.2
Phosphoenol-pyruvate
Glyoxylate and dicarboxylate metabolism

B


METHANE METABOLISM
Methane
CoM-S-S-CoB
Formaldehyde
Ribulose-P pathway
D-Xylulose-5P
D-Ribulose-5P
D-Arabino-Hex-3-ulose-6P
D-Fructose-6P
Methanogenesis
Coenzyme B
Coenzyme M
Methanol
Methyl-CoM
MtaA
MtaB
MtaC
MttA
MttB
MttC
MtbA
MtbB
MtbC
MtmB
MtmC
Trimethylamine metabolism
Trimethylamine-N-oxide
Trimethylamine
Dimethylamine
Methylamine
N-Methyl-L-glutamate
Folate biosynthesis
Pentose phosphate pathway
Xylulose-P pathway
D-Glyceraldehyde-3P
Glycerone
Glycerone-P
D-Fructose-1,6-P2
Sulfur metabolism
Glycine, serine and threonine metabolism
L-Serine
3-Phosphoserine
3-Phosphonooxypyruvate
Glycerate-3P
Tetrahydrofolate
Carbon fixation
5,10-Methylene-tetrahydrofolate
Glycine
Serine pathway
S-Hydroxymethylglutathione
Formyl-MFR
N5-Formyl-THMPT
5,10-Methenyl-THMPT
5,10-Methylene-THMPT
5-Methyl-THM(S)PT
THM(S)PT
Methanofuran (MFR)
Coenzyme F420H2
F420
F420H2
Glyoxylate
Hydroxypyruvate
D-Glycerate
Glycerate-2P
Glyoxylate and dicarboxylate metabolism
S-Formylglutathione
Malyl-CoA
L-Malate
Oxaloacetate
Phosphoenolpyruvate
Acetyl-CoA pathway
Formate
CO2
CO
ACDS
Acetyl-CoA
Pyruvate
Acetate
Acetylphosphate
Nitrogen metabolism
Coenzyme B biosynthesis

c


METHANE METABOLISM
Methane
CoM-S-S-CoB
Formaldehyde
Ribulose-P pathway
Methanogenesis
Coenzyme B
Coenzyme M
Methanol
Methyl-CoM
Trimethylamine metabolism
Trimethylamine-N-oxide
Trimethylamine
Dimethylamine
Methylamine
N-Methyl-L-glutamate
Folate biosynthesis
Pentose phosphate pathway
Xylulose-P pathway
D-Glyceraldehyde-3P
Glycerone
Glycerone-P
D-Fructose-1,6-P2
Sulfur metabolism
Glycine, serine and threonine metabolism
L-Serine
3-Phosphoserine
3-Phosphonooxypyruvate
Glycerate-3P
Tetrahydrofolate
Carbon fixation
Glycine
Serine pathway
Formyl-MFR
THM(S)PT
Methanofuran (MFR)
Glyoxylate
Hydroxypyruvate
D-Glycerate
Glycerate-2P
Glyoxylate and dicarboxylate metabolism
Malyl-CoA
L-Malate
Oxaloacetate
Phosphoenolpyruvate
Acetyl-CoA pathway
Formate
CO2
CO
ACDS
Acetyl-CoA
Pyruvate
Acetate
Acetylphosphate
Nitrogen metabolism
Coenzyme B biosynthesis

D

Predicted methane metabolism in GAL by KEEG. The four methanogenesis pathways through: the conversion of CO_2 to methane [A], through the conversion of acetate to methane [B], the conversion of methanol to methane [C] and the conversion of methylamine/dimethylamine/trimethylamine to methane [D].

4.3.5. Mapping of nitrogen and sulfur pathways in DAL and BLA

The pathways mapped for nitrogen and sulfur metabolisms based on SEED annotation of metagenomes of DAL and BLA, were complete or near complete. In nitrogen pathway mapping of DAL and BLA, ‘nitrogen fixation’, ‘assimilatory nitrate reduction’, ‘dissimilatory nitrate reduction’ and ‘denitrification’ subsystems were predicted (Figs. 22 and 23). All the predicted genes in subsystems of nitrogen metabolism are directly or indirectly involved in the ammonification process in both DAL and BLA.

Important genes for a complete denitrification process were predicted in the BLA metagenome. As for the DAL metagenome, all genes for denitrification except the final nitric-oxide (NO) reductase expression regulator (NosR) were predicted, which led to incomplete mapping of denitrification. However, net nitrification was absent in both sites. In general, different types of enzymes, transporter proteins and transcription factors involving in nitrogen metabolism were all predicted from these data.

For sulfur metabolism, more than 100 thousand reads were annotated in DAL and BLA (Appendix 7). The average percentage of assigned reads to the subsystems is similar in both sites, except for the subsystems At5g37530 and Galactosylceramide and sulfatide metabolism. The KEGG mapping of sulfur metabolism for DAL and BLA depicted complete pathways, where sulfate is converted to sulfide (‘assimilatory sulfate reduction’ pathway) and sulfate is converted to thiosulfate and back to sulfate (‘SOX system’) (Fig. 24 A and B).

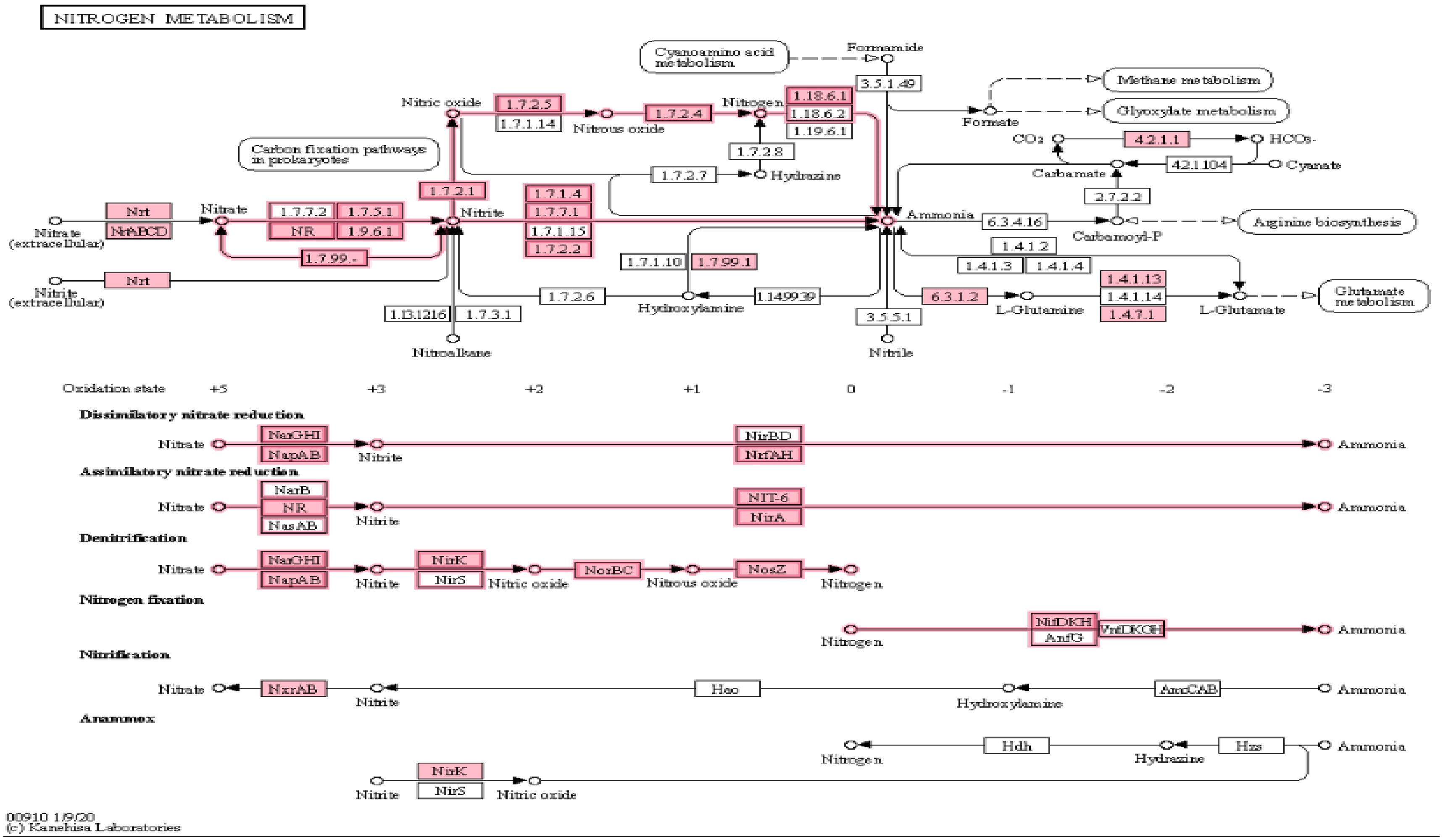


Fig. 22 Predicted nitrogen cycle in BLA

Predicted nitrogen cycle in BLA based on SEED-KEEG metabolic pathway mapping. Pink highlight color represents the presence of the specific enzymes and/or proteins for the specific metabolic reaction and red line indicated the metabolic pathway in the nutrient cycle.

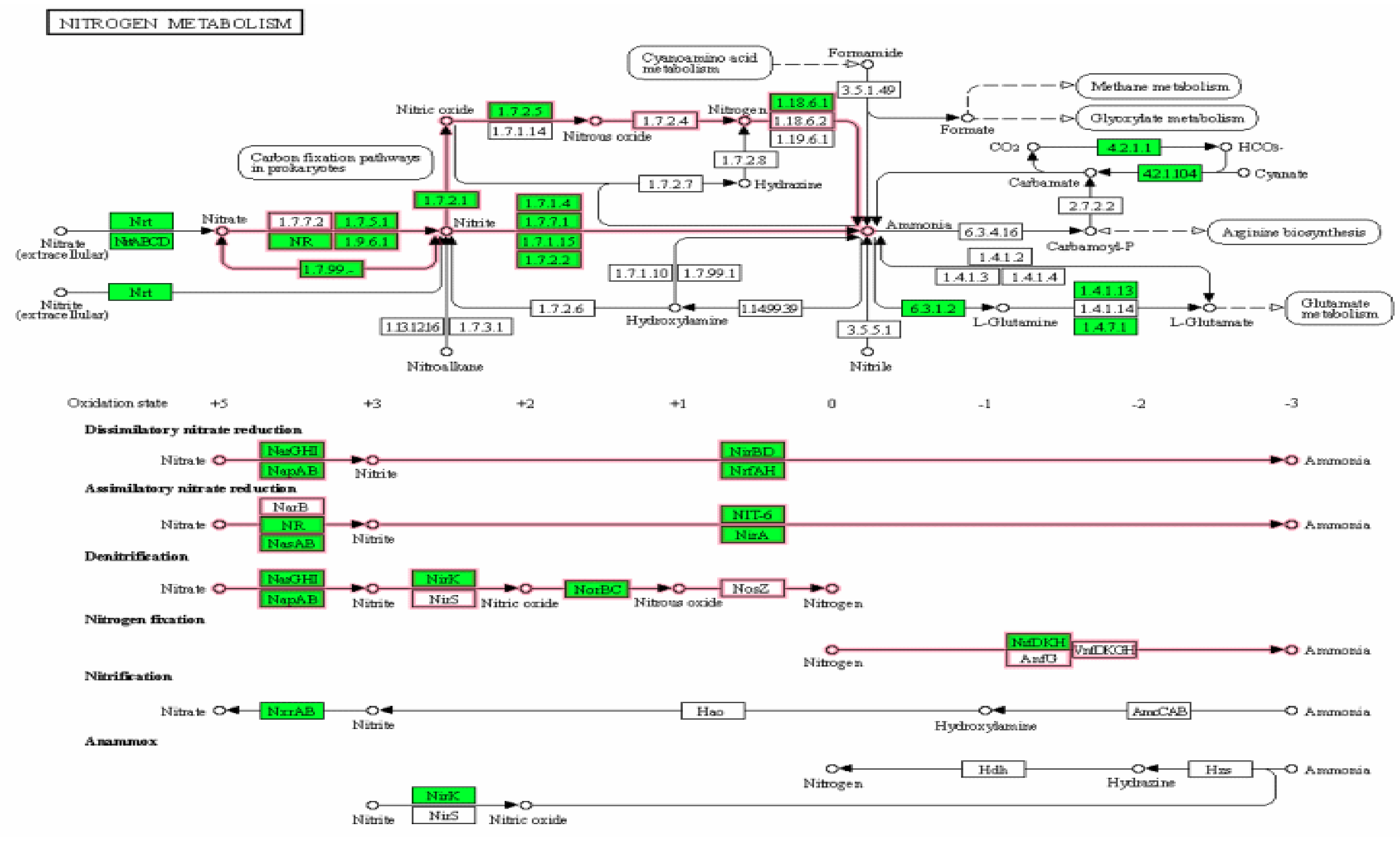
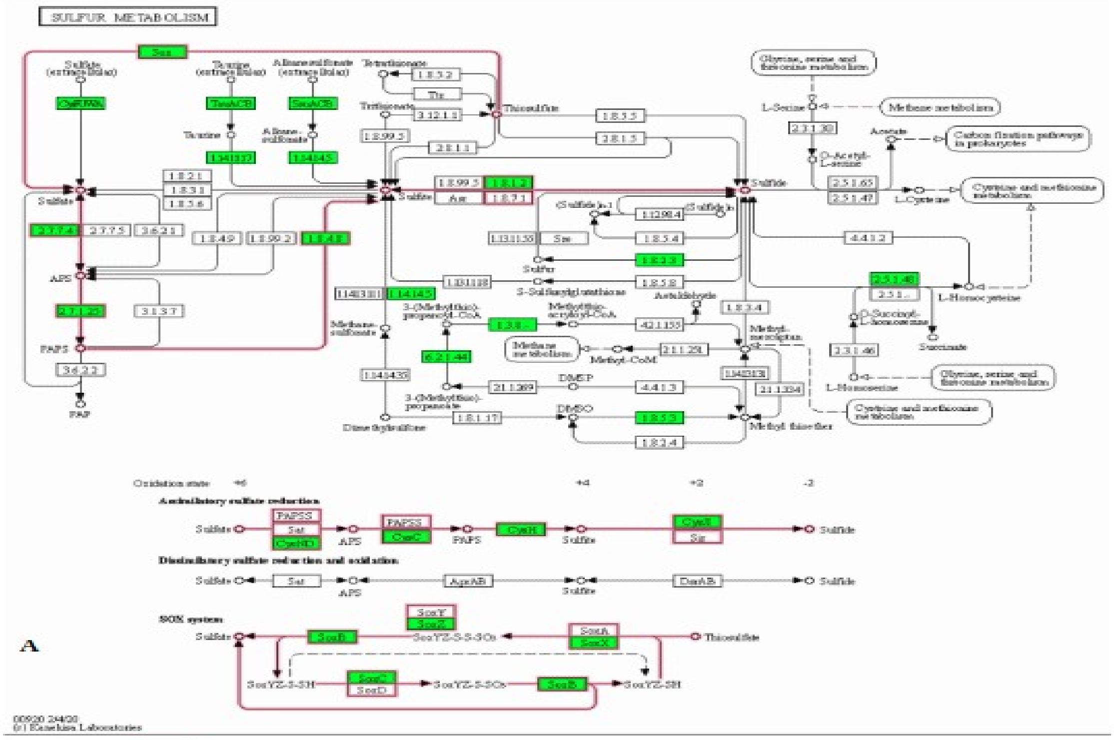
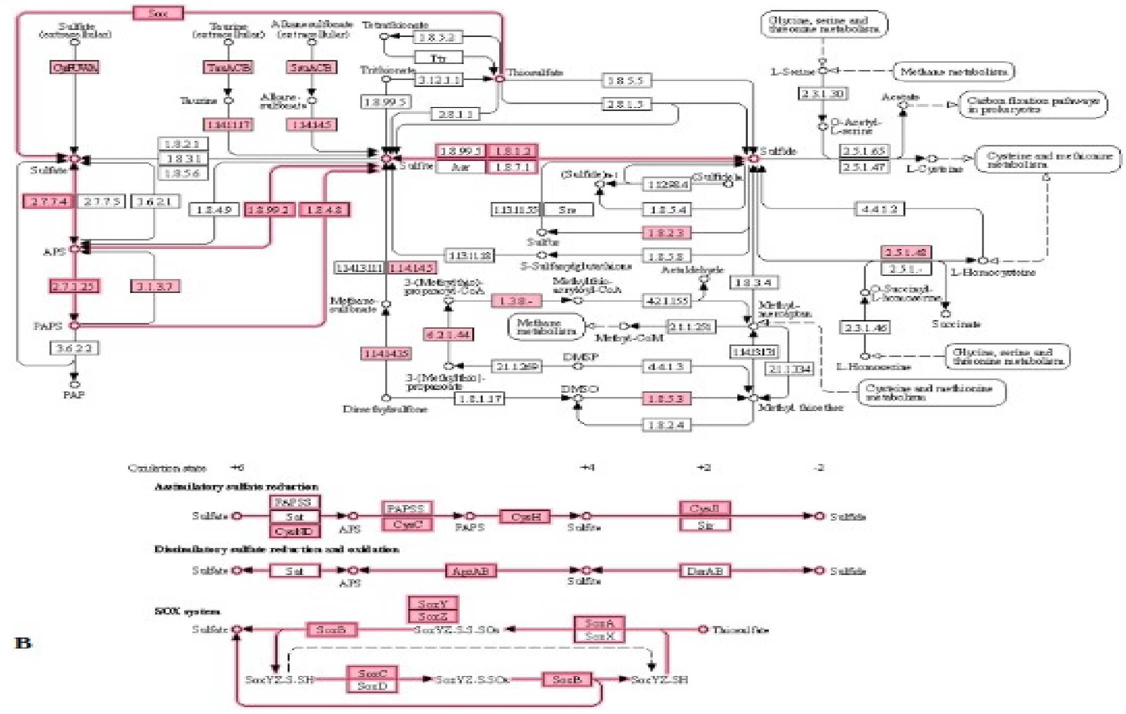


Fig. 23 Predicted nitrogen cycle in DAL

Predicted nitrogen cycle in DAL based on SEED-KEEG metabolic pathway mapping. Green highlight color represents the presence of the specific enzymes and/or proteins for the specific metabolic reaction and red line indicated the metabolic pathway in the nutrient cycle.


SULFUR METABOLISM
Sulfate (extracellular)
Taurine (extracellular)
Alkanesulfonate (extracellular)
Tetrathionate
Trithionate
Thiosulfate
Sulfate
APS
PAPS
PAP
Sulfite
Sulfide
Sulfur
L-Serine
O-Acetyl-L-serine
L-Cysteine
L-Homocysteine
O-Succinyl-L-homoserine
Succinate
L-Homoserine
Acetate
Acetaldehyde
Methanethiol
Methyl-mercaptan
Dimethyl sulfide
DMSO
DMSP
Dimethylsulfone
Methanesulfonate
3-(Methylthio)-propanoate
3-(Methylthio)-propanoyl-CoA
Methylthio-acryloyl-CoA
Methyl-CoM
S-Sulfanylglutathione
Glycine, serine and threonine metabolism
Methane metabolism
Carbon fixation pathways in prokaryotes
Cysteine and methionine metabolism
Oxidation state
+6
+4
+2
-2
Assimilatory sulfate reduction
PAPSS
Sat
CysND
CysC
CysH
CysJI
Sir
Dissimilatory sulfate reduction and oxidation
AprAB
DsrAB
SOX system
SoxY
SoxZ
SoxA
SoxX
SoxB
SoxC
SoxD
SoxYZ-S-SH
SoxYZ-S-SO3
SoxYZ-SH
00920 2/4/20
(c) Kanehisa Laboratories

A


Sulfate (extracellular)
Taurine (extracellular)
Alkanesulfonate (extracellular)
Tetrathionate
Trithionate
Thiosulfate
Sulfate
APS
PAPS
PAP
Sulfite
Sulfide
Sulfur
L-Serine
O-Acetyl-L-serine
L-Cysteine
L-Homocysteine
O-Succinyl-L-homoserine
Succinate
L-Homoserine
Acetate
Acetaldehyde
Methyl-mercaptan
Dimethyl sulfide
DMSO
DMSP
Dimethylsulfone
Methanesulfonate
3-(Methylthio)-propanoate
3-(Methylthio)-propanoyl-CoA
Methylthio-acryloyl-CoA
Methyl-CoM
S-Sulfanylglutathione
Glycine, serine and threonine metabolism
Methane metabolism
Carbon fixation pathways in prokaryotes
Cysteine and methionine metabolism
Oxidation state
+6
+4
+2
-2
Assimilatory sulfate reduction
PAPSS
Sat
CysND
CysC
CysH
CysJI
Sir
Dissimilatory sulfate reduction and oxidation
AprAB
DsrAB
SOX system
SoxY
SoxZ
SoxA
SoxX
SoxB
SoxC
SoxD
SoxYZ-S-SH
SoxYZ-S-SO3
SoxYZ-SH

B

Sulfur cycle in DAL (A) and BLA (B) based on SEED-KEGG. Green and pink highlight colors represent the predictions of the specific enzymes and/or proteins for the specific metabolic reaction and red line indicated the metabolic pathway in the nutrient cycle.

4.3.6. Prediction of 1º, 2º Structure and 3D modeling of selected nitrogen fixing proteins in DAL and BLA

Complete or partial translated ORFs, with different percentages of identity to homologous proteins involved in nitrogen metabolisms were predicted. The MEGAN-Aligner aligned and connected two or more contigs to predict complete or near complete ORFs. In the case of nitrogen metabolism, relatively numerous contigs with predicted genes were binned to the classes of “ammonia assimilation” and “nitrate /nitrite ammonification”. Several contigs were also binned into the “nitrogen fixing” class of nitrogen metabolism in DAL and BLA. Functional orthology was searched and matched for 92% and 81% of the predicted ORFs involved in nitrogen metabolism for DAL and BLA, respectively.

From DAL assembled metagenome, the translation products of 2 complete or near complete ORFs (homologous to [4Fe-4S] ferredoxin and cysteine desulfurase) and 8 partial ORFs (homologous to nifK, nifH, nifB, nifX, nifZ, nifQ, nifV, nifE) were predicted. On the other hand, from BLA, translation products of 5 complete or near complete ORFs (homologous to 4Fe-4S ferredoxin and cysteine desulfurase, NifB, NifT and NifW) and 8 partial ORFs (homologous to NifD, NifK, NifA, NifH, NifX, NifE, NifN and NifV) were predicted. The results of prediction of the primary and secondary structures of the translated products of the complete or near complete ORFs of query proteins versus reference proteins using two different servers are listed in Table 11.

Table 11 Comparison of predicted secondary (2^0) structures

Site	Contig’s node number	Amino acid length	Predicted protein	GOR4				SOPMA			
				α helix (%)	Extended sheet **(%)**	**β turn (%)**	Random coil **(%)**	**α helix (%)**	Extended sheet **(%)**	**β turn (%)**	Random coil **(%)**
BLA	Node_4344	551	Cysteine desulfurase	38.40	10.53	0	51.00	34.30	13.79	8.17	43.74
Reference [a]	WP_149425408176_609	534	Cysteine desulfurase	41.01	9.18	0	49.81	33.90	12.92	6.93	46.25
DAL	Nodes_5315_35911	393	Cysteine desulfurase	44.78	13.74	0	41.48	40.97	17.30	8.91	32.82
Reference [a]	WP_149425408176_609	393	Cysteine desulfurase	46.56	10.69	0	42.75	42.75	16.79	8.14	32.82
BLA	Node_27591	72	4Fe-4S Ferredoxin	0	19.44	0	80.56	2.78	29.17	6.94	61.11
DAL	Node_15557	72	4Fe-4S Ferredoxin	0	19.44	0	80.56	2.78	29.17	6.94	61.11
Reference [b]	WP_063195668.1:	72	4Fe-4S Ferredoxin	0	27.78	0	72.22	1.39	26.39	5.56	66.67
BLA	Node_7493_18684	457	NifB	36.54	16.63	0	46.83	40.70	16.41	7.22	35.67
Reference [c]	WP_091676580	457	NifB	31.29	19.04	0	49.67	43.11	14.88	7.00	35.01
BLA	Node_62099	69	NifT	14.49	31.88	0	53.62	11.59	31.88	13.04	43.48
Reference [d]	WP_137042982	69	NifT	14.49	31.88	0	53.62	8.70	36.23	7.25	47.83
BLA	Node_21959	109	NifW	57.80	5.50	0	36.70	62.39	7.34	1.83	28.44
Reference e	WP_016918478.1	109	NifW	53.21	5.50	0	41.28	66.06	6.42	3.67	23.85

[a]= reference sequence used for aligning ORF predicted for cysteine desulfurase
[b]= reference sequence used for aligning ORF predicted for 4Fe-4S Ferredoxin
[c]= reference sequence used for aligning ORF predicted for NifB
[d]= reference sequence used for aligning ORF predicted for NifT
[e]= reference sequence used for aligning ORF predicted for NifW

Table 12 Sequence similarity of query proteins to the top three reference proteins

Site	Node number and length of assembled translated sequence (amino acid)		**1st Reference**				**2nd Reference**				**3rd Reference**			
			Accession number	Max score	Identity (%)	E value	Accession number	Max score	Identity (%)	E value	Accession number	Max score	Identity (%)	E value
BLA	Node_4344	551	WP_149425408.1	753	69.41	0.0	PKO82724.1	753	67.94	0.0	WP_152089440.1	731	67.44	0.0
	Node_27591	72	WP_063195668.1	127	83.33	$9e^{-37}$	WP_080136056.1	126	84.72	$2e^{-36}$	WP_100178681.1	126	84.72	$2e^{-36}$
	Node_7493_18684*	457	WP_091676580.1	870	88.8	0.0	WP_137042984.1	858	87.53	0.0	CCJ08068.1	858	88.33	0.0
	Node_62099	69	WP_137042982.1	133	91.3	$4e^{-39}$	OYW42097.1	127	86.96	$7e^{-37}$	WP_014892121.1	127	85.51	$1e^{-36}$
	Node_21959	109	WP_016918478.1	150	65.14	$1e^{-44}$	WP_109027636.1	144	62.39	$2e^{-42}$	WP_018267935.1	143	61.47	$4e^{-42}$
	Nodes_5315_35911*	393	WP_149425408.1	697	84.22	0.0	WP_109039789.1	677	81.93	0.0	WP_152089440.1	677	81.93	0.0
	Node_15557	72	WP_063195668.1	127	83.33	$9e^{-37}$	WP_080136056.1	126	84.72	$2e^{-36}$	WP_100178681.1	126	84.72	$2e^{-36}$
DAL														

* Merged nodes that aligned to the same reference sequences

The primary sequence analyses for all predicted proteins in both BLA and DAL showed a difference in composition of amino acids to their respective reference proteins. The number of negatively charged amino acids was predicted to be higher than the number of positively charged amino acids in all selected proteins as in the case of their respective reference proteins. The results of secondary structure predictions obtained from Gor4 and SOPMA were different for all query protein sequences (Table 11). While Gor4 was only able to predict the three basic secondary structural elements (helix, sheet and coil), SOPMA further predicted beta turns and the results are shown in both numeric and graphical forms. The difference among the secondary structure prediction results produced by the respective servers could probably result from the use of different indices in making the prediction by the servers.

The amino acid sequence identity of the selected predicted proteins was checked against the reference proteins. The maximum percentage of similarity obtained was for the nitrogen fixing protein “NifT” from BLA at 91.3% identity to *Pseudolabrys sp. FHR47* (WP_137042982.1). The rest of selected predicted protein sequences have less than 80% identity to their respective reference proteins Table 12. The difference among the secondary structure prediction results produced by the respective servers could probably result from the use of different indices in making the prediction by the servers. The 3-D models of the two commonly found query proteins and their respective references are shown in Figures 25 and 26.

Table 13 Predicted protein motifs, patterns and profiles

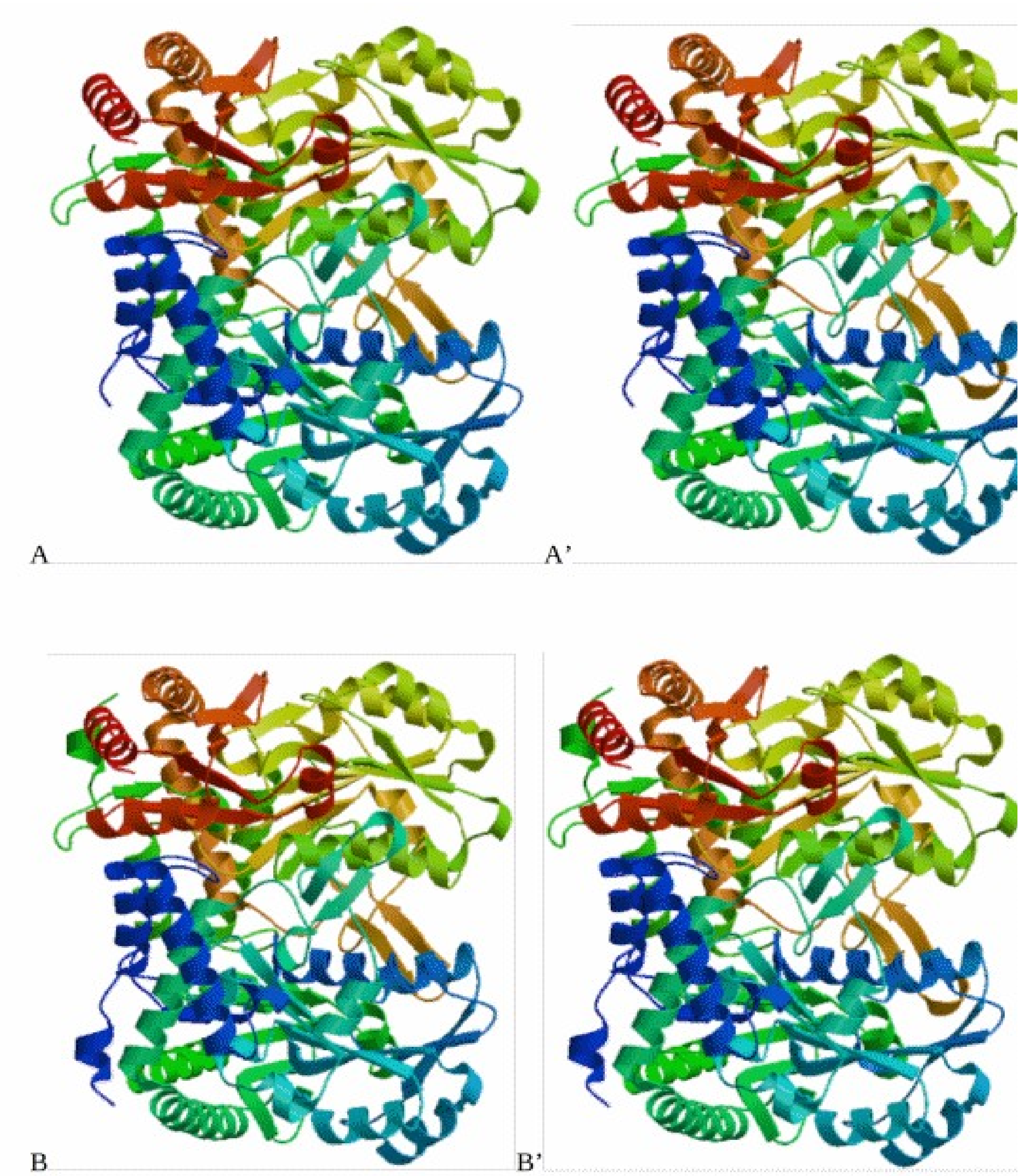
Sampling site	Contig's node number and protein coding region	Overlapping protein entries and regions	Biological function	Binding sites[a] -molecular function	
BLA	Node_4344 (147-547)	IPR015424 PyrdxlP-dep_Trfase: 4-403 IIPR000192 Aminotransferase class V: 26-395 IPR015422 PyrdxlP-dep_Trfase_dom1: 9-398 IPR015421 PyrdxlP-dep_Trfase_major: 37-297 IPR010970 Cys_dSase_SufS: 8-403	cysteine metabolic process (GO:0006534)	K [227] /catalytic residue (GO:0003824) A, T [95-96], I [99], H [124], D [201], A, Q [203-204], S [224], H, K [226-227]-Pyridoxal 5'-phosphate binding pocket (GO:0030170)	
DAL	Nodes_5315_35911(1-393)	IIPR000192 Aminotransferase class V: 17-386 IPR015421 PyrdxlP-dep_Trfase_major: 28-288 IPR010970 Cys_dSase_SufS: 1-392 IPR015422 PyrdxlP-dep_Trfase_dom1: 1-389 IPR015424 PyrdxlP-dep_Trfase: 1-392	cysteine metabolic process (GO:0006534)	K [218]/catalytic residue (GO:0003824) A, T [86-87], I [90], H [115], D [192], A, Q [194-195], S [215], H, K [217-218]]- Pyridoxal 5'-phosphate binding pocket (GO:0030170)	
BLA	Node_27591(1-72)	SSF54862 4Fe-4S Ferredoxins: 4Fe-4S_Fe-S-bd	NA	NA	NA
DAL	Node_15557 (1-72)	SSF54862 4Fe-4S Ferredoxins: 4Fe-4S_Fe-S-bd	NA	NA	NA
BLA	Node_7493_18684 (35-434)	IPR005980 Nitrogenase cofactor biosynthesis protein NifB: rSAM	coenzyme biosynthetic process (GO:0009108)	C [67], I [69], C [71], Y, C, N [73−75], L [115], G, D [121−122], S, P, H [149−151], T [173], V [228], N, I [254−255]	catalytic activity (GO:0003824), iron-sulfur cluster binding (GO:0051536), metal ion binding (GO:0046872),4 iron, 4 sulfur cluster binding (GO:0051539)
BLA	Node_62099 (1-69)	NifT/FixU (IPR009727) NA	nitrogen fixation (GO:0009399)	NA	NA
BLA	Node_21959 (1-109)	Nitrogen fixation protein NifW (IPR004893) NA	nitrogen fixation (GO:0009399)	NA	NA

[a] Positions of binding sites on assembled contigs with the corresponding residues

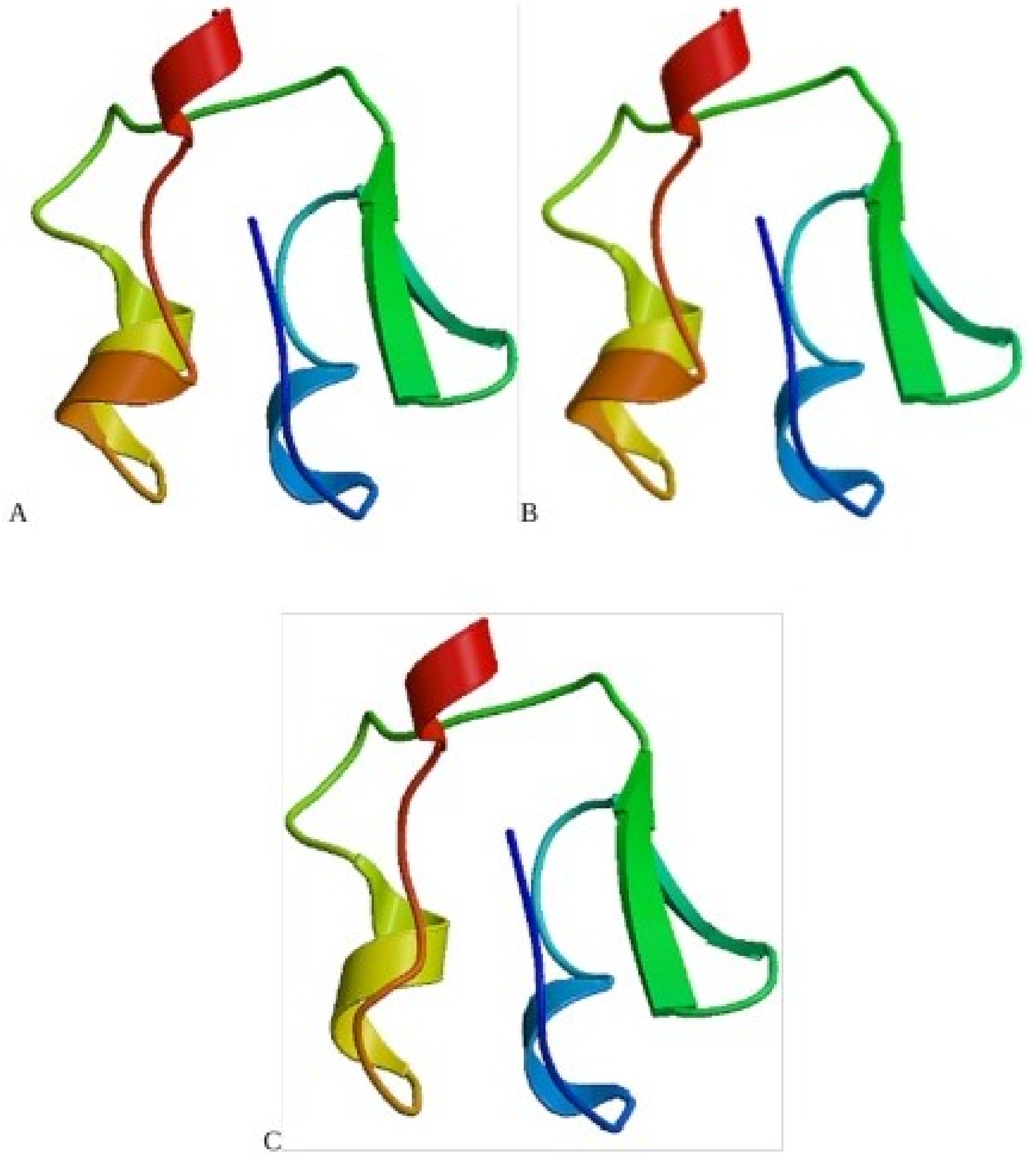
Table 14 Swiss model program parameters and Ramachandran plot output

		Protein model and PDB ID	Target and Template Identity (%)	Ramachandran plot distribution of residue		
				Most Favoured (%)	Allowed (%)	Outliner (%)
BLA	Node_4344_cysteine_desulfurase	Cysteine desulfurase from *Legionella pneumophila Philadelphia* 1(6c9e.1.A)	53.88	95.72	0.76% [A463 VAL, B463 VAL, B199 HIS, A199 HIS, A464 PRO, B464 PRO]	0.32% [A385 GLU, B480 SER]
	WP_149425408.1:76-609 SufS family cysteine desulfurase	Cysteine desulfurase from *Legionella pneumophila Philadelphia 1*(6c9e.1.A)	51.13	95.47	0.88% [A448 VAL, B448 VAL, B184 HIS, A184 HIS, A513 PHE, A449 PRO, B449 PRO]	0.63% [A370 GLU, B465 SER, A419 LEU, B419 LEU]
	Node_27591_ferredoxine	2[4FE-4S] ferredoxin from *Pseudomonas aeruginosa* (2fgo.1)	48.39	88.71	0.0%	0.0%
	WP_063195668.1:1-72 ferredoxin	2[4FE-4S] ferredoxin from *Pseudomonas aeruginosa (2fgo.1)*	53.23	88.71	1.61% [A23 VAL]	3.51% [A33 VAL, A62 VAL]
DAL	Nodes_5315_35911_cysteine_desulfurase	Cysteine desulfurase from *Legionella pneumophila Philadelphia 1(6c9e.1.A)*	54.96	95.65	0.77% [A311 VAL, B311 VAL, B47 HIS, A47 HIS, A312 PRO, B312 PRO]	0.32% [A233 GLU, B328 SER]
	WP_149425408.1:213-605 SufS family cysteine desulfuras	Cysteine desulfurase from *Legionella pneumophila Philadelphia 1*(6c9e*.1.A)*	51.15	96.55	0.64% [A47 HIS, B47 HIS, B311 VAL, A312 PRO, B312 PRO]	0.97% [B213 VAL, A213 VAL, A233 GLU, B328 SER, A282 LEU, B282 LEU]
	Node_15557_ferredoxine	2[4FE-4S] ferredoxin from *Pseudomonas aeruginosa (2fgo.1)*	48.39	88.71	0.0%	0.0%
	WP_063195668.1:1-72 ferredoxin	2[4FE-4S] ferredoxin from *Pseudomonas aeruginosa (2fgo.1)*	53.23	88.71	1.61% [A23 VAL]	3.51% [A33 VAL, A62 VAL]

According to prediction outputs, the SWISS-MODEL used cysteine desulfurase from *Legionella pneumophila Philadelphia 1* and 2[4Fe-4S] ferredoxin from *Pseudomonas aeruginosa* as templates. In all the models predicted (Figs. 25 and 26), the structures are in agreement with the results of the secondary structure earlier predicted for the query as well as the reference sequences (Waterhouse *et al.*, 2018). Alpha helix and random coils dominated the structure of the predicted cysteine desulfurase while random coils and extended sheets dominated the structure of the predicted [4Fe-4S] ferredoxin. Functional motifs within the cysteine desulfurase and NifB protein families were identified (Table 13). The Pyridoxal 5'-phosphate binding motif within the cysteine desulfurase and FeS/SAM binding motif with in FeMo cofactor biosynthesis protein (nifB-like) were determined. The predicted positions of conserved and active sites of the query proteins with the corresponding amino acid residues are listed in Table 14.



3D modeling of cysteine desulfurase from DAL [A] with reference protein WP_149425408.1 [A’] and cysteine desulfurase from BLA [B] with reference homologous protein WP_149425408.1 [B’] using SWISS-MODEL Workspace (Waterhouse *et a*l., 2018).



3D modeling of [4Fe-4S] ferredoxin from DAL [A], BLA[B] and reference homologous protein WP_0631956681172 [C] using SWISS-MODEL Workspace (Waterhouse *et al.*, 2018).

4.4. Discussion

4.4.1. Taxonomic signature of assembled sequences

MetaSPAdes and MEGAN are among the most reliable and fastest tools for assembling highly uneven metagenomic reads and taxonomic binning, respectively (Meiser *et al.*, 2017). Half of the metagenome sequences in this study are covered by contigs length ranging between larger than or equal the N50 contig size reported on Table 7. The taxonomic assignment of assembled contigs by MEGAN was in consensus with the Metaphlan2 mapping done in previous chapter. Large number of assembled contigs from DAL, GAL and BLA were binned to bacteria, primarily to the phylum Proteobacteria, as in the case of most extreme acidic environments studied so far (Spear *et al.* 2006; Johnson *et al.*, 2001; Méndez-García 2015; Mesa *et al.,* 2017; Schuler *et al.* 2017). On the other side, numerous assembled contigs from LA, LADP5, LADP6 and MUP were binned to Archaea with an almost exclusive assignment to the phylum Halobacteria.

For better understanding of microbial ecology and metabolism, justification of functional potential of the extremophiles through metagenomic sequence data was advantageous since WGS increased the rate of detection and prediction of microbial diversity and genes (Schmeisser *et al.*. 2007; Su 2012; Ranjan *et al.*, 2015; Abreu 2018; Tran and Phan, 2020). Analyzing pathways of formation of complex polymers from simpler elements or monomers is important to know the stress response processes of microbial mediated nutrient cycling in extreme environments such as hot springs, hypersaline lakes and volcanic pools (Li *et al.* 2018).

Insilico investigation of gene functions in the current study was performed only on less than 33.33% of the total metagenomes of all sampling sites. Yet, the four important biogeochemical cycles for sustaining life were identified and numerous protein families involving in the nutrient biocycling were annotated in the HsAP than in HsAqS. Carbon and nitrogen are the main constituents of living organisms as they are the essential components for all organic polymers and the building blocks of genetic materials (Widdison and Burt 2008). Thus, large number of contigs with translated ORFs encoding enzymes, transcription factors and different proteins for the carbon metabolism were predicted in this study. Moreover, few but essential genes for converting different forms of biologically important nutrients such as phosphorous, sulfur and nitrogen were also projected.

4.4.2. Carbon fixation in HsAP and HsAqS

Reports showed that carbon dioxide (CO_2) is the most oxidized state of carbon which makes it a stable and inert molecule (Schuchmann and Müller, 2013; Liu *et al.*, 2020^{b}; Lemaire *et al.*, 2020). Nowadays, carbon fixation by photosynthetic organisms is an Eco-friendly biological carbon sequestration process that has captured the interest of researchers to reduce the negative effect of climate changes (Schuchmann and Müller, 2013; Liu *et al.*, 2020^{b}). There are different autotrophic carbon fixation pathways existing among bacteria, archaea and eukaryotes and the two-step process where CO_2 is branched on a reactive group (carboxylation) and then reduced is the common setting (Berg *et al.*, 2010; Tikh and Schmidt-Dannert, 2013; Ferrari and Freschi, 2019). In our study, contigs with ORF for protein families involved in carbon fixation were among the many predicated in all the sampling sites except for GAL. Only four enzymes (transketolase [EC:2.2.1.1], ribulose-bisphosphate carboxylase large chain [EC:4.1.1.39], ribulose-phosphate 3-epimerase [EC:5.1.3.1] and ribose 5-phosphate isomerase A [EC:5.3.1.6]) involved in reductive pentose phosphate cycle (Calvin cycle) were mapped from GAL.

Reductive pathway acetyl-CoA (also known as the Wood–Ljungdahl Pathway) is the condensation of two one-carbon units (carbon dioxide) to form a two-carbon compound (acetate) under an anaerobic condition (Ragsdale and Pierce, 2008). In this pathway, two molecules of CO_2 are reduced in parallel to methyl radical and carbon monoxide (CO) and form acetyl-CoA (Ljungdahl, 1986; Ragsdale and Pierce, 2008). The enzyme CO dehydrogenase/acetyl-CoA synthase (EC:1.2.7.4, 1.2.99.2 and 2.3.1.169) receives the methyl group and combines it with the CO and releases coenzyme A to form an acetyl-CoA molecule. In all brine metagenomes of this study, this key enzyme and formate dehydrogenase (EC:1.2.1.43) were not detected. Instead, acetate kinase (EC:2.7.2.1) and phosphate acetyltransferase (EC:2.3.1.8), that catalyze the conversion of acetate into acetyl-CoA were detected in DAL, BLA and LADP6. Normally, the action of these two important enzymes is observed in aceticlastic methanogens for a “reverse” Wood–Ljungdahl pathway and catalyze the conversion of acetyl-CoA into acetate to generate metabolic energy (ATP) (Ragsdale and Pierce, 2008; Lemaire et al., 2020).

The other important carbon fixating pathway is the reductive Tri Carboxylic Acid cycle (rTCA), where energy rich carbon compounds are produced from carbon dioxide and water (Katiyar *et al.*, 2018; McCully *et al.*, 2019; Prakasham and Kumar, 2019). The rTCA is a reverse of the oxidative TCA (also known as krebs cycle or citric acid cycle), that is used to produce energy by all aerobic organisms through oxidation of pyruvate (Katiyar *et al.*, 2018; McCully *et al.*, 2019; Prakasham and Kumar, 2019). To reverse the reactions of the oxidative citric acid cycle, the three important enzymes succinate dehydrogenase, NAD^+-dependent 2-oxoglutarate dehydrogenase and citrate synthase must be replaced by succinate dehydrogenase/ fumarate reductase, ferredoxin-dependent 2-oxoglutarat synthase and ATP citrate lyase, respectively. The key enzyme for the rTCA pathway is succinate dehydrogenase iron-sulfur protein/ fumarate reductase (EC 1.3.99.1), which catalyzes the reduction of fumarate to succinate using ubiquinol as an electron donor (Cecchini *et al.*, 2002). We observed that the metagenomes from the studied cites (except for GAL) encoding two distinct succinate dehydrogenase/fumarate reductase (EC:1.3.99.1) namely, flavoprotein subunit (K00239) and iron-sulfur protein (K00240) (Appendix 10). Furthermore, an anaerobic enzyme citrate lyase subunit beta (citE; K01644; EC:4.1.3.6) which catalyzes the division of citrate to oxaloacetate and acetyl-CoA was detected (Katiyar *et al.*, 2018; McCully *et al.*, 2019; Prakasham and Kumar, 2019). Another important enzyme of the rTCA cycle is “isocitrate dehydrogenase or IDH” that catalyzes the reversible conversion of isocitrate to 2-oxoglutarate in prokaryotes (Wang *et al.*, 2013). However, 2-oxoglutarate ferredoxin:oxidoreductase (EC:1.2.7.3) which catalyzes the reductive carboxylation of succinyl-CoA to 2-oxoglutarat was not detected.

In our study, the main pathway mapped by KEGG is the carbon fixation associated to photosynthetic organisms (ko00710) and two complete cycles (Reductive pentose phosphate cycle or Calvin cycle and Crassulacean acid metabolism CAM) were identified. Carbon fixation is an integral part of photosynthesis, especially in plants (Tikh and Schmidt-Dannert, 2013; Cummins *et al.*, 2018). For natural photosynthetic microorganism, the combination of light energy conversion and CO_2 fixation has reduced the dependency of the organism on organic material as a carbon source (Tikh and Schmidt-Dannert, 2013). Ribulose 1,5-bisphosphate carboxylase-oxygenase (RuBisCO) is the key enzyme responsible for fixing atmospheric CO_2 within the Calvin-Benson cycle and carboxysomes are polyhedral inclusion bodies that contain the enzyme (Berg et al., 2010; Tsai and Yeates, 2011; Tikh and Schmidt-Dannert, 2013; Hobot, 2015; Cummins et al., 2018). The first extremophile, where intact carboxysomes were first isolated from *Halothiobacillus neapolitanus* (Dou *et al.*, 2008; Tsai and Yeates, 2011). Two important enzymes (carbonic anhydrase (EC:4.2.1.1) and ribulose bisphosphate carboxylase (EC:4.1.1.39)) were identified mainly in DAL, GAL and BLA. Two reactions are catalyzed by Ribulose bisphosphate carboxylase (RuBisCO); carboxylation of D-ribulose 1,5-bisphosphate (CO_2 fixation) and oxidative fragmentation of the pentose substrate (photorespiration). However, RuBisCO has low affinity for CO_2, though CO_2 concentrating mechanism (CCM), Carbonic anhydrase converts bicarbonate inside the carboxysome to CO_2 (Dou *et al.*, 2008; Cummins *et al.*, 2018; Ferraroni, 2019). Yet in some extremophiles (like in metagenomes of Lake As'ale and MUP), where carbonic anhydrase gene was lacking, it is suggested that the carboxysomes acts as a diffusional barrier for CO_2 (Dou *et al.*, 2008; Tsai and Yeates, 2011).

The other two variants of photosynthetic carbon assimilation (C4 photosynthesis and CAM) are considered as carbon concentrating mechanisms under dark conditions and form adjuncts of the Calvin cycle (Raghavendra *et al.,* 2017; Ferrari and Freschi, 2019). Both C4 photosynthesis and CAM are phenomenon occurring in plants (Raghavendra *et al.,* 2017; Ferrari and Freschi, 2019). However, in recent studies on metagenomes of extreme environments indicated that autotrophic CO_2 fixation through CAM was one of the major pathways representing carbon fixation (Wilkins *et al.,* 2019; Meziti *et al.,* 2020).

The predicted CO_2 fixation strategies in DAL, BLA, MUP, and coastal parts of Lake As'ale are important information for carbon capture and storage. The application of autotrophic organisms in industries to avoid or minimize CO_2 emission was not as efficient as needed because of low overall solar-to-product energy conversion efficiencies; hence, mixed-substrate conversions techniques were considered as an option (Santero *et al.,* 2016; Steiger *et al.,* 2017). In this study, the conversion of CO_2 into biomass is predicated to be carried out by groups of different bacterial and archaeal families in their respective communities. The understanding of syntrophic interactions and its contribution in the CO_2 fixation pathways in the studied extreme environments can be used as an insight into design environmental biotechnological applications under hypersaline and different pH conditions to minimize CO_2 emission from industries.

4.4.3. Methanogenesis in GAL

Methanogenesis pathways were exclusively predicted in GAL, while carbon fixation was absent. Excess emission of CO_2 in GAL has been reported by Master, (2016). Methanogenesis is the final step in anaerobic degradation of carbon and it is the major contributor of greenhouse gases and the global carbon cycle (Berghuis *et al.*, 2019; Kurth *et al.,* 2020; Lemaire *et al.,* 2020). Methanogenesis and sulfate reduction coexist, where microbially-generated high concentrations of methane and hydrogen sulfide have been observed in different environments (Mitterer, 2010; Sela-Adler *et al.*, 2017; Sorokin *et al.*, 2017; Carr *et al.*, 2018; McGenity and Sorokin, 2019). Though different dissimilatory sulfite reductases were predicted in GAL (Appendix 7), the depleted concentration of SO_4^{-2} either previously recorded in Gebresilassie *et al*. (2011) or measured in this study (Table 1) indicated the presence of thermodynamically favorable condition for methanogenesis (Mitterer, 2010; Sela-Adler *et al.*, 2017; Sorokin *et al.*, 2017; Carr *et al.*, 2018; McGenity and Sorokin, 2019).

There are three methanogenesis pathways (hydrogenotrophic, methylotrophic, and acetoclastic), where the enzyme methyl-coenzyme M reductase (Mcr) is the common catalyst (Berghuis *et al.*, 2019). Acetoclastic (acetate-utilizing) and methylotrophic methanogenesis pathways are active in freshwater sediments (or anaerobic digestors) and marine sediments (or other anoxic systems rich in methylated substrates), respectively (Berghuis *et al.*, 2019).

Several enzymes important for hydrogenotrophic and acetoclastic methanogenesis have been predicted from the metagenome data. Hydrogenotrophic methanogenesis is one of the most primitive of extant metabolisms of respiration where the organisms that use it can grow autotrophically by using H_2 as electron donor and CO_2 as sole carbon source and electron acceptor (Boyd *et al.*, 2020; Kurth *et al.*, 2020; Berghuis *et al.*, 2019; Tilahun *et al.*, 2021[b]). In hydrogenotrophic methanogenesis (Fig. 20B), energy is conserved by coupling to reductive acetyl-CoA or Wood–Ljungdahl pathway using methyl-tetrahydromethanopterin: coenzyme M methyltransferase (Mtr). The function of Mtr is to transfers the methyl group from the Wood–Ljungdahl pathway to coenzyme M, which is subsequently converted to methane by methyl coenzyme M reductase and using coenzyme B as an electron donor. Two subunits of Mtr (subunits G and H) and six subunits of methyl coenzyme M reductase (alpha I and II, beta I and II, gamma I and II) involved in hydrogenotrophic methanogenesis have been exclusively predicted in GAL metagenome.

Like the hydrogenotrophic methanogenesis, acetoclastic methanogenesis also include many of the proteins involved in Wood–Ljungdahl pathway, but acetyl-CoA synthetase needs to be present to activate acetate to acetyl-CoA (Carr *et al.*, 2018; Han *et al.*, 2019; Berghuis *et al.*, 2019). Yet acetyl-CoA synthetase enzyme for acetoclastic methanogenesis was not predicted in GAL metagenome. However, acetate kinase (EC 2.7.2.1) and phosphate acetyltransferase (EC 2.3.1.8) were predicted in GAL metagenome. While acetate kinase is predicted to convert acetate to acetyl phosphate, phosphate acetyltransferase converts the acetyl phosphate to acetyl-CoA, which will be further converted to methane (Fig. 20C). These two enzymes are common in anaerobic bacteria species and are important for anaerobic decomposition of complex organic matter to methane (Groshong and Blevins, 2014; Han *et al.,* 2019; Tilahun *et al.,* 2021[b]).

Among the three pathways of methanogenesis, methylotrophic methanogenesis was not mapped for GAL because important enzymes such as methane monooxygenase (MMO) were not detected in the metagenome. Methylotrophic methanogenesis is believed to be independently evolved ancient pathway where methylated C1 compounds, such as methanol are first activated by specific methyltransferase and then reduced to methane (Vanwonterghem *et al.,* 2016; Tilahun *et al.,* 2021[b]; Wang *et al.,* 2021).

Most methanogens grow optimally under neutral conditions while some methanogens found in natural environment are halophilic, acidophilic or acid-tolerant (Braüer *et al.*, 2006; Sanz *et al.*, 2011; Sorokin *et al.*, 2017; Carr *et al.,* 2018; McGenity and Sorokin, 2019; Han *et al.,* 2019). Normally, methanogens were grouped to class I and class II of Euryarchaeota in the domain of archaea (Braüer *et al.,* 2006; Sanz *et al.,* 2011; Carr *et al*., 2018; Han *et al.*, 2019; Berghuis *et al.*, 2019). But as in the case of GAL, the Proteobacteria and Cytophaga, Fusobacterium, and Bacteroides (CFB) bacteria have also been previously detected in methane-rich marine and lake sediments (Dong *et al.,* 2006; Mitterer, 2010; Sela-Adler *et al.,* 2017).

Nearly all essential genes important for methanogenesis were detected in GAL metagenome and were binned to Protebacteria (Appendix 11). There are basically two important principles about essential genes: (1) they are monogenic and more conserved than nonessential genes and (2) essential genes required in vitro are more conserved than those in vivo signifying the different strategies of organisms under stressful conditions (Grazziotin *et al.*, 2015; Tilahun *et al.*, 2021[b]). Most of GAL's predicted genes annotated for methanogenesis are binned to alpha and beta Proteobacteria. Functional analysis using orthologous protein groups showed over prediction of methyl coenzyme M reductase genes. Among the many the predicted essential genes for methanogenesis in GAL metagenome, genes that have more than 95% coverage are; dimethylamine methyltransferase corrinoid protein, dimethylamine:corrinoid methyltransferase, formylmethanofuran dehydrogenase (molybdenum) operon gene G, methyl coenzyme M reductase I alpha subunit and monomethylamine methyltransferase corrinoid protein (Appendix 11).

In methanogenic bacteria CO_2 is anaerobically convert to methane using different types of coenzymes, such as Coenzyme M. Usually, methanogens first generate formate and CO by reduction of CO_2, which are subsequently fixed to transfer the carbon toward their central metabolism where methane is the end product (or acetate for acetogens) (Lemaire *et al.*, 2020). The formyl-methanofuran dehydrogenase (Fwd) complex is the key enzyme that catalyze the reduction of CO_2 and the conversion of formate ($HCOO^–$) into a formyl group (Lemaire *et al.*, 2020). Both isoforms of this enzyme containing either a molybdenum- or tungsten (formylmethanofuran dehydrogenase (molybdenum) subunit C (EC:1.2.99.5) and formylmethanofuran dehydrogenase (tungsten) subunit D (EC:1.2.99.5)) were predicted in GAL.

The percentage similarities of the predicted essential genes for methanogenesis in GAL range from 52% (with 95% query coverage) to 100% (with 18-100% query coverage) and mostly binned to Bradyrhizobiaceae and Burkholderiacaea. Archaeal families are the major players in methanogenesis and Bradyrhizobiaceae and Burkholderiacaea are the least referred for this pathway (Brauer *et al*., 2006; Kimura *et al.*, 2010; Sanz *et al.*, 2011; Carr e*t al.*, 2018; McGenity and Sorokin, 2019; Han *et al*., 2019). However, it is well known that most bacteria carry horizontally transferred essential genes that are widespread in archaea and vice versa (Kimura *et al.*, 2010; Grazziotin *et al.*, 2015). Different strategies are deployed when organisms such as Bradyrhizobiaceae and Burkholderiacaea are exposed to or stressed by methane-rich environments. Cooperative (syntrophic) interactions among bacteria is an important survival strategy of microbial consortia in different environments (Mee *et al.*, 2014; Fowler *et al.*, 2016). This metabolic cross feeding among microbes enables cooperative growth and metabolic exchange within a shared pool of micronutrients (Mee *et al.*, 2014; Fowler *et al.*, 2016; Kurade *et al.*, 2019). In methanogenesis, cooperation of fermentative bacteria to methanogens in basis of the transfer of hydrogen, formate, or acetate to make the degradation of electron-rich substrates thermodynamically favorable was described (Sieber *et al.*, 2010).

Overall, the prediction of methanogenesis in GAL present a great biotechnological opportunity for large scale methanogen production as acidophilic extremophiles are extremely important in anaerobic decomposition of sewage, industrial waste management and as a methane-renewable energy source (Braüer *et al.*, 2006; Kimura *et al.*, 2010; Sanz *et al.*, 2011; Grazziotin *et al.*, 2015). The understanding of syntrophic interactions and its contribution in methanogenesis pathway in GAL can be used as an insight into design industrial biotechnological applications under hypersaline and acidic conditions for the production of methane bio-gas.

4.4.4. Nitrogen and Sulfur metabolisms in DAL and BLA

The nitrogen cycle is the second most important nutrient cycle to organisms because it provides the building blocks of DNA and proteins (Widdison and Burt, 2008). The contigs with protein coding genes for nitrogen metabolism were predicted in both the DAL and BLA metagenomes. In the case of sulfur metabolism, assimilatory sulfate reduction (ASR) and sulfate oxidation pathways were predicted and mapped (Appendix 7 and Fig. 22). Just like nitrogen, sulfur in living organisms needs to be reduced in order to form proteins and other macromolecules (Peck, 1961; Schiff and Fankhauser, 1981; Kawano *et al.,* 2018).

Assimilatory sulfate reduction (ASR) is a pathway used by prokaryotes and eukaryote to convert inorganic sulfate to sulfide, which is further incorporated into carbon skeletons of amino acids to form Cysteine (Cys) or homo Cysteine (homo-Cys) (Kushkevych *et al.,*2020). The ASR in DAL and BLA is supplied by the Sox pathway found in many sulfur bacteria oxidizes thiosulfate to sulfate (Grabarczyk and Berk, 2017). Different types of L-cysteine S-thiosulfotransferase were predicted for DAL (SoxA) and for BLA (SoxX), while S-disulfanyl-L-cysteine oxidoreductase (SoxD) was absent in both DAL and BLA. These types of thiosulfate utilization by the Sox system typically found in Proteobacteria (Gregersen *et al.,* 2011).

In both DAL and BLA, the ASR pathway was facilitated by the sulfate adenylyl transferase/ adenylyl-sulfate kinase (Cys C/ Cys N), phosphoadenosine phosphosulfate reductase (Cys H), sulfite reductase (NADPH) flavoprotein alpha-component (Cys J) enzymes. Organisms that follow ASR pathway usually have the ability to grow on sulfate as their sole source of sulfur. According to Yuan *et al.,* 2021, in aerobic conditions, assimilatory sulfate reduction is promoted while nitrate is reduced to nitrite, hence indicating that some microorganisms can change their energy conservation and survival strategy through respiration control. The H_2S produced in ASR is usually not detectable but is a transient intermediate for the biosynthesis of cysteine, hence ASR pathway doesn't produce toxic hydrogen sulfide (Kushkevych *et al.*,2020). Vital steps in the transformation and movement of nitrogen in the two environments were mapped based on the translated ORFs present in the metagenome. The predicted paths in the nitrogen cycle of the two acid ponds showed that the main biochemical reaction is reduction of nitrogen. The distinct absence of a nitrogen oxidation pathway and the prediction of allantoin utilization can be indicative of the strict nitrogen budget in the two extreme environments and the necessity for the residing extremophiles to maintain the bioavailability of nitrogen (Vogels and Drift, 1976).

Further deduction of metabolism of ammonia to L-glutamate also showed that glutamate is the main precursor which contributes amine groups for the biosynthesis of amino acids, nucleic acid bases, and aminated carbohydrates in both sampling sites (Kim and Gadd, 2008; Meng *et al.,* 2016). Inhibition of nitrification in hypersaline environments suppresses denitrification and anaerobic ammonium oxidation (anammox), resulting in efficient ammonium recycling and also enriches ammonium in surface brine due to the release of 15N-depleted NH_3 gas (Isaji, 2019). Proteins involved in nitrification were not predicted in this study. Partial and/or complete ORFs for all important orthologous proteins involved in denitrification were predicted in the metagenomes of both ponds. But in the DAL metagenome, crucial protein coding genes for the final reduction of NO to N_2 were missing, causing incomplete mapping of denitrification. Still, it is not uncommon among extremophiles, especially thermophiles to undergo incomplete denitrification (Chen *et al.* 2002; Hedlund *et al.* 2011). On the other hand, protein coding genes, homologous to copper containing nitrite reductase (NirK) and nitrite reductase associated c-type cytochorome (NirN) were only predicted in the DAL metagenome. The NirN has been proven to be an essential cofactor for the cytochrome associated nitrite reductase NirS, which is very important for catalyzing the reduction of NO_2^- to NO in denitrification process (Adamczack 2014; Ward 2001).

One of the most important functional annotations performed for the metagenome reads from both acid ponds is the deduction of nitrogen fixation. In principle, the prospect of genetic manipulation of plants to fix nitrogen independently and fulfill the energy demand has prompted researchers to explore different genetic resources and methods (Dixon, 1997; Ivleva, 2016; Li, 2016; Burén *et al.,* 2018). Fixing atmospheric nitrogen by the nitrogenase enzyme is a characteristic present only in diazotrophs and it influences the amount of nutrient present within an ecosystem (Hamilton *et al.,* 2011; Jones, 2016; Messer, 2016; Lesser *et al.,* 2018). In DAL and BLA, the genes involved in nitrogen fixation were predicted only from Proteobacteria OTUs, particularly, from Alphaproteobacteria. As in the case of many diazotrophic bacteria, either in poly-extreme environment or not, the predicted nitrogenase from the metagenomes of DAL and BLA has two components; the NifH-[4Fe-4S] cluster and nitrogen fixation (nif) genes for the biosynthesis of Mo–Fe cofactor (Dixon, 1997; McGlynn, 2013).

Molecular and biochemical data are limited to construct the best model for biosynthesis of nitrogenase for all heterotrophic free-living nitrogen fixing bacteria (Lesser *et al.,* 2018). Still, a common understanding exists on how the environment determines the quantity and organization of nitrogen fixing genes in different diazotrophs (Lesser *et al.,* 2018; Pedersen, 2018; Pérez, 2017). The broadly accepted minimal gene cluster for Mo-dependent nitrogenase biosynthesis consists of nifDKH and nifBEN sets (McGlynn, 2013; Pérez, 2017; Black and Santos, 2015; Wang, 2013). However, the predicted sets of nif genes in the two metagenomes are quite different from each other even if both sample sites are hyper saline and highly acidic. Differences among the types of accessory proteins involved in maturation of nitrogenase and trafficking molybdenum to nitrogenase was also observed in the two acid ponds. For instance, partial ORFs of nifQ and nifZ were predicted only in DAL. The NifQ and NifZ proteins are known to be involved in trafficking molybdenum to nitrogenase and maturation of the P-clusters contained within the Mo–Fe protein, respectively (Imperial *et al.,* 1984; Hernandez *et al.,* 2008; Meng, 2016; Isaji, 2019; Chen *et al.,* 2002; Hedlund *et al.,* 2011; Jimenez-Vicente *et al.,* 2018; Jimenez-Vicente *et al.,* 2019).

On the other hand, complete ORFs of nifT and nifW and partial ORFs of nifA and nifN genes were predicted only in the BLA. The NifW protein is involved in the maturation processes of the NifD protein, whereas NifA is an enhancer-binding protein that binds to specific DNA sequences upstream of nif genes and NifT has some suppressive effects on nitrogenase (Morett *et al.,* 1991; Nonaka *et al.,* 2019). Both blastx and blastp results showed that many of the predicted Nif proteins in both acid ponds are highly divergent with less than 85% average similarity of translated ORF sequences to orthologs of Nif proteins. As a principle to sequence analysis, more distant species will have more variable sequences of orthologous proteins (Koonin and Galperin, 2003). Thus, divergent groups of free-living nitrogen fixing bacteria may inhabit these two extreme environments of the Danakil Depression.

Further structural examination on translated ORFs for cysteine desulfurase and ferredoxin revealed that the predicted hypothetical proteins from both acid ponds are quite different from their respective orthologous reference proteins (Table 12). Cysteine desulfurase (NifS/SufS), a homodimeric enzyme, mobilizes sulfur from cysteine via a pyridoxal 5′-phosphate (PLP) dependent mechanism and is known to be contributing in many biosynthesis reactions (Zheng *et al.,* 1993; Black and Santos, 2015; Poudel *et al.,* 2018; Kiyasu, 2000). If NifS is absent, the activities of both nitrogenase component proteins can be negatively impacted (Zheng *et al.,* 1993; Poudel *et al.,* 2018). Scrutiny of the pairwise amino acid sequence alignments with the reference protein along with the analysis of the corresponding structural models showed that functionally important residues in the predicted enzymes were conserved. In particular, the lysine residue with catalytic activity and other residues interacting as binding sites for the Pyridoxal 5′-phosphate are identically conserved in both predicted proteins as well as the reference (Table 11).

The [4Fe-4S] ferredoxins predicted from both DAL and BLA have identical ORFs and the closest homologous protein is the ferredoxin of *Bradyrhizobium sp. AT1*(WP_063195668.1) with only 83.0% amino acid sequence identity. This protein is an important, but not the primary, electron donor for dinitrogenase reductase (Poudel *et al.,* 2018; Egener *et al.,* 2001). Yet, it is known to have an effect on the rapid "switch-of" of nitrogenase activity in response to ammonium (Zheng *et al.,* 1993; Egener *et al.,* 2001). The nodes with ORFs, representing query proteins cysteine desulfurase and ferredoxin from DAL have more than 80.0% identity with conserved active sites to the reference proteins, while 69.41% and 83.3% similarity was observed between the query proteins from BLA and their respective reference proteins (Table 12). Normally, amino acid sequence similarity less than 80.0% is more difficult to reliably infer similar function (Pearson, 2013).

However, according to the secondary and tertiary structure analyses, all the important residues on the sites of major twists and turns as well as the active sites are present to maintain the structural and functional integrity of the predicted proteins (Tables 11, 13). Hence, regardless of the apparent dissimilarity of primary amino acid sequences of the query proteins to the reference proteins, the similarities of the secondary and 3-D structures of the predicted cysteine desulfurase and ferredoxin from DAL and BLA to the reference proteins clearly indicated functional orthology (Pearson, 2013) (Table 14). The BLA metagenome provided us with many complete ORFs to further study possible novel Nif proteins. The predicted proteins (NifB, NifT and NifW) amino acid sequences were dissimilar to the closest homologous proteins (Tables 11). Just like in the case of cysteine desulfurase, functionally important residues in the predicted NifB were conserved (Table 11)

4.5. Conclusion

Reconstruction of genomes that exist within a metagenomic sample was an important computational step of insilico analyses. Four important biogeochemical cycles for sustaining life were predicted. As carbon is the main constituent of living organisms and the essential component for all organic polymers, large number of contigs with translated ORFs encoding enzymes, transcription factors and different proteins for the carbon metabolism were predicted. While carbon fixation was predicted for DAL, BLA, Lake As’ale and MUP, methanogenesis was exclusively predicted in GAL. Different types of enzymes, transporter proteins and transcription factors involving in nitrogen, carbon fixation and sulfur metabolisms were predicted from all studied metagenomic data. Yet not all genes important for the above-mentioned metabolisms were equally predicted in all sampling sites. In some sampling sites, complete pathways were predicted while in others, some nutrients metabolic pathways were completely or partially absent. The predicted genes and metabolic pathways from all sampling sites indicated that nutrients metabolism is carried out by groups of different bacterial and archaeal families within their respective communities enabling cooperative growth and metabolic exchange within a shared pool of micronutrients.

Chapter Five

5. Visualization of Dallol Extremophiles using Transmission Electron Microscopy and Evaluation of Extremophilic Growth Media

Abstract

In the present work, we describe microscopy approach that can improve direct visualization and studies of microbes in Dallol acid pond coded as DAL-P4 (equivalent to DAL in previous chapters). Different shapes and sizes of microbes were identified using Transmission Electron Microscopy (TEM). Cellular components such as ribosomes, pili, glycoprotein cell wall, capsule and many nano-sized extracellular membrane vesicles (EMV) were identified. The EMVs were originated from outer membrane of the cells and were observed protectively aggregating on some actively dividing cells. Viability and resistance of cells preserved in "sucrose-lysis buffer" to the detrimental conditions of lysing and extracting buffers was investigated using specially concocted liquid growth media at different pH conditions. Growth rates were calculated using Optical Density (OD) measurements of microbial inoculum in culture media and higher rate was recorded for acidic condition than for basic condition. Furthermore, removal of phenolic complex was indirectly observed while growth curve of inoculum in basic media with $FeCl_3$ was plotted. In general, this preliminary investigation indicates great biotechnological and bioremediation potentials of DAL-P4 extremophile in the fields of medicine, different industries and environment.

Keywords; Extremophiles, EMV, Transmission Electron Microscope

5.1. Introduction

Living things developed several types of cellular or molecular mechanisms which are considered as responses to environmental changes (Boor, 2006; Meena *et al.,* 2017; Guan *et al.,* 2017; Mozaheb and Mingeot-Leclercq, 2020). Environmental stresses usually have significant influence on microbial physiology which helps them to develop a series of strategies to resist the imposed stress (Boor, 2006; Meena *et al.,* 2017; Guan *et al.,* 2017; Mozaheb and Mingeot-Leclercq, 2020). Modifications on structure and composition of the cell membrane, adjustment of permeability and activities of transporters are some examples of mechanisms used by microbes to alleviate the destructive effects of stressors (Guan *et al.,* 2017). On the other hand, many microbes (Gram positive, Gram negative and Archaea) are reported to produce nano-sized membrane vesicles to adapt or survive disastrous environmental changes (Toyofuku *et al.,* 2019; Byvalov *et al.,* 2020; Mozaheb and Mingeot-Leclercq, 2020).

Normally, membrane vesicles are diverse in their origin, components, and size but all are membrane derived (Byvalov *et al.,* 2020; Mozaheb and Mingeot-Leclercq, 2020). Based on their origin (plasma membranes and endosomes), membrane-derived vesicles and exosomes are the two types of known extracellular vesicles (Raposo and Stoorvogel, 2013; Turturici *et al.,* 2014). Size can also be used to distinguish EMV as nano-sized or micron-sized vesicles (Momen-Heravi *et al.,* 2013; Yokoyama *et al.,* 2017; Hosseinkhani *et al.,* 2017). Most archaea, Gram-positive and Gram-negative bacteria can produce and discharge EMV that contain specific cargo molecules (Toyofuku *et al.,* 2019; Byvalov *et al.,* 2020; Mozaheb and Mingeot-Leclercq, 2020). Some of the previously identified molecules that have been transported by EMVs are: virulence factors, DNA, antibiotics and eukaryotic host defence factors, cell signaling factors, and so on (Toyofuku *et al.*, 2019; Byvalov *et al.*, 2020).

The extremophiles of Dallol acid pools should be physiologically well equipped to systematically adapt and survive the unsteady extreme conditions. Preliminary metagenomics study conducted in 2015 on DAL, showed the existence of many biotechnological important extremophiles, with predicted different types of genes involved in nutrient cycles (Tilahun *et al.*, 2021[a]). Among the main discoveries about Dallol extremophiles, their resistance to keep the integrity of cell membrane during stressful conditions (cell lysing conditions for DNA extraction) is the important and intriguing one (Tilahun *et al.*, 2021[a]). Hence further investigations were needed to understand the coping mechanisms of extremophiles in DAL. In addition, viability and propagation of extremophiles in standard chemical cell lysing buffers needed to be checked. Therefore, the objective of this investigation was to study microscopical cellular adaptation mechanisms of the extremophilic cells and to characterized the microbial growth in specialized media that are known to have lysing effect on normal cells.

5.2. Materials and methods

5.2.1. Sample collection

Sample collection was done on Dallol active spring at the crate that was located at UTM 0639983 Easting/1574575 and previously studied in 2015. 2 litters of brine samples were collected using sterile High-Density Polyethylene (HDPE) bottles and 50 ml conical plastic tubes (BD Falcon TM, Greiner Bio-One, Germany) for cell harvesting and physicochemical analysis. The samples were stored in an icebox containing cooling elements and transported immediately to a temporary working space at Yara Dallol BV, Dallol Camp for processing. Cell pellets were collected after spinning using Hettich Mikro 20 Centrifuge at 10,000 rpm for 5 min using 1.5 mL Eppendorf safe-lock tubes. Sterile 25% NaCl solution was used to wash and pool the cell pellets. Total of 175 mL brine sample was centrifuged to pellet microbial cells and collected into 8 x 1.5 mL Eppendorf safe-lock tubes. Half of the pelleted cells were preserved using sterilized sucrose lysis buffer (0.75 M sucrose, 20 mM EDTA, 200 mM NaCl, and 50 mM Tris-HCl, pH 9.0) and the other half were preserved in absolute ethanol (96%) and stored in iceboxes containing cooling elements and then stored at -20^{0}C at AAU until further transported to Max Planck Institute Munster, Germany.

5.2.2. Physico-chemical measurements

Onsite measurement of pH was done using 0-14 pH indicator paper sticks (Fisherbrand™) while *insitu* measurements of temperature and conductivity were avoided. Salinity was measured using a refractometer (DIGIT-0120 ATC, VWR, UK) by diluting the brine samples 1 to 10X with deionized water. Comprehensive physical and chemical analysis of the brine sample was conducted at Environmental Chemistry Laboratory, Ethiopian Public Health Institute (under the serial number 936757). Iron (Fe^{3+}) was measured at Center of Environmental Science, AAU.

5.2.3. Cell imaging using transmission electron microscopy

Cell suspension from the preserving buffers were taken and processed for Transmission Electron Microscopy (TEM) at Max Planck Institue, Munster Germany. Fixation and preparation for TEM were according to Mildner and Zeuschner, (2017). During preparation, 1µl preserved samples were fixed in 2% paraformaldehyde, 2% glutaraldehyde in 0.1 M cacodylate buffer and allowed to settle overnight on a poly-Llysin-coated cover slip. The fixed samples become post fixed with osmium tetra oxide followed by dehydration in ethanol. The last dehydration step and the resin mixtures were done in propylene oxide where the cover slips were transferred into a glass dish. the initially added resin was polymerized at 60°C for at least 6 hours. Pure gelatin capsules without bottom are placed over the cells after the final filling with epon. The sample blocks can then be removed by putting the whole cover slip repeatedly in liquid nitrogen and hot water. Small sectioning of the flat epon embedded samples was performed, as cells were directly located at the edge of the block. The embedded sample in epon was counter-stained using undiluted uranyl acetate (Ura-ace) and lead (Pb). About 30% sucrose was used as a negative control and 120 kV TEM Hitachi HU-12A, Japan was used for imaging.

5.2.4. Preparation of liquid media for Dallol brine samples isolates

Special liquid media were formulated to culture Dallol microbes based on the extraordinary survival track records of cells under lysing conditions. The components of the specialized liquid media are adopted from the compositions of sucrose lysis buffer (Mitchell and Takacs-Vesbach, 2008) and phenol-DNA/RNA extraction techniques. The pH for these two specialized culture media was adjusted to be acidic (pH 2) and and basic (pH 9). For the acidic culture media, the list of components is; 1 M NaCl, 20 mM Na_2EDTA, 0.75 M sucrose, 7% phenol solution (saturated with 0.1 M citrate buffer, pH 4.3 ± 0.2, BioReagent), 3% ethanol (96%), 5.2 mM $FeCl_3$, 18 mM $MgCl_2.6H_2O$, 0.8% nutrient broth and 100ml of distilled H_2O. All components of the liquid media except phenol, ethanol and $FeCl_3$ were sterilized together using autoclave while the rest were filter sterilized separately. The autoclave and filter sterilized components were mixed once the former cooled off. For the second type of media, all the components and formulation procedure listed for the acidic medium was repeated except the phenol solution used was equilibrated with 10 mM Tris HCl, pH 8.0, 1 mM EDTA, BioReagent. Further adjustment of pH was done for the acidic medium (final pH=2) and the basic medium (final pH=9) using 3 M HCl and 1M NaOH, respectively.

The two types of special media were inoculated with thawed 0.01% (v/v) cell suspension preserved in sucrose lysis buffer (stored at -20 ^{0}C). All culture samples were prepared in triplet with corresponding controls and were incubated at 55^{0}C. Optical density (OD) was measured using UV/VIS spectrophotometer and polystyrene cuvettes. Depending upon the nature of the phenol, the phenol-$FeCl_3$ complex can have different intensities of color. Hence, in our investigation, the acidic culture media had yellowish color while the basic culture had violet color. Therefore, the wavelength was set to 600nm for the acidic media and 380nm for the basic media because the Phenol-Fe (III) complex is deep violet in color. Indirect measurement of phenol-Fe (III) removal rate was obtained by calculating OD at 380 nm relative to the abiotic control used as blank. Growth rate of microbes in the two types of media was determined in terms of generation per hour and doubling time, using the equations k=log(Xt)−log(X0) x 1/0.301t (equation 1) and Doubling time (tgen)=1/k (equation 2). Growth curve was plotted using measured optical density (OD) over time (T).

Further microscopy was done to take images of cells in the special culture media. From basic culture media, 15 mL of samples were centrifuged at 10,000g for 15 min at 4 °C. Cell pellets were washed three times with Phosphate-buffered saline (PBS) where excess supernatant was aspirated before re-suspending in pre-warmed “Cell Tracker TM Green CMFDA” working solution according to the manual. Cell Tracker solution was removed after centrifugation and the cells were dispensed onto slide for microscopy. Image was analyzed at absorption and fluorescence emission maxima of 492 and 517 (nm), respectively using Zeiss Axio Observer. None fluorescent image was analyzed using Zeiss Axio Observer with Objective Plan-Apochromat 100x/1.40 Oil DIC M27 under oil immersion. All images were taken using AxioCamMR3 camera at Max Planck Institute of Molecular and Biomedicine, Munster, Germany.

5.3. Results

5.3.1 Water chemistry of DAL-P4

The hydro-chemical analysis result showed that the salinity of the acid pond studied was 40% (Table 15). The amount of some ions such as Ca^{2+}, Mg^{2+} and PO_4^- could not be determined due to the complex nature of the sample and presence of interference for the analysis (according to the report from EPHI).

The pH values obtained indicated a highly acidic environment. High contents of NH_4^+, NO_3^- and NO_2^- (343.5 mg/L, 1.7 g/L and 3.46 mg/L, respectively) was measured, while zero or nearly zero amount of HCO_3^- and CO_3^- were obtained (Table 15). Additional cations were also measured Na+, Fe (III)+ and K+ with the amount 92.41 g/L, 1.35 g/L and 14.99 g/L ,respectively recorded.

Table 15 Physico-chemical analysis of DAL-P4 brine sample

Parameter	DAL-P4
Turbidity: (NTU)	0
TDS (g/L)	551
Electrical Conductivity at 25ºC (µS/cm)	803000
PH at 25ºC	0***
Carbonate alkalinity as $CaCo_3$ (g/L)	Nil *
Bicarbonate alkalinity as $CaCo_3$ (g/L)	Nil *
Total hardness as $CaCo_3$ (g/L)	48
Silica, total (SiO_2) (g/L)	Na**
NH_4^+ (mg/L)	343.5
Na^+ (g/L)	92.41
K^+ (g/L)	14.99
Ca^{2+} (g/L)	NA**
Mg^{2+} (g/L)	NA**
Cl^- (g/L)	187.44
NO_2^- (mg/L)	3.46
NO_3^- (g/L)	1.76
F^- (g/L)	0.31
HCO_3^- (g/L)	Nil*
CO_3^- (g/L)	Nil *
PO_4^- (ortho) (g/L)	NA**
Salinity	0.4
Fe^{3+} (g/L)	1.35

'Nil*' zero or nearly zero

'NA**' unable to determine the concentration of some elements due to the complex nature of the sample and presence of interferences

'***' negative pH value is the actual value obtained following standard analysis procedure. It is not error in laboratory analysis or instrumental problems

5.3.2. Transmission Electron Microscopy of Extremophiles

During Transmission Electron microscopy, the staining uranyl acetate and lead (Pb) helped to increase the contrast under the electron beam that pass through the very thin sections of a sample. This double contrasting with uranyl acetate and lead citrate produced the highest electron density and image contrast. As a result, production of numerous Extracellular Membrane Vesicles (EMVs) was observed from samples preserved in ethanol and sucrose lysis buffers. The TEM of first sectioning showed differently sized and shaped aggregates of electron-dense EMVs (Fig. 27-30). The estimated size of the EMVs is consistent in both samples and is less than 30 nm in diameter (Fig. 27 E and Fig. 28 E-F).

The TEM of the second section from cells preserved in absolute ethanol and sucrose lysis buffers helped to identify important cellular components. Morphological diversity of microbes associated with the presence of different internal and external cellular structures were revealed by TEM. In addition, the presence of membrane-bound vesicles originating from the microbial outer membranes were observed (Fig. 28 D and 29 D-G and). These electron-dense extrusions from the cell's surfaces are morphological indicatives of secretory activities into the environment.

The observed microbes in absolute ethanol buffer were different in size (diameters varying from 200 nm to 500 nm) and with typical structures such as, cytoplasmic membrane with cell wall and capsule (Fig. 29 B-C [3-5]) or without capsule (Fig. 29 D-G) and with distinct cytoplasmic electron-density differences (Fig. 29 G [9]). Large variations in the thickness of the microbial cell envelope, capsule and external fibril proteins (pili) were also observed (Fig. 29 B-C). On the other side, for microbes preserved in sucrose lysis buffer (Fig. 30), cellular activities such as cell division (Fig. 30 B [1] and E [7]) and cellular structures such as nucleoid (Fig. 30 E [6]), ribosomes (Fig. 30 D [5]) cytoplasmic membrane with cell wall and capsule (Fig. 30 D [4]), membrane proteins (Fig. 30 C [3]) and nano fiber production by EMV (Fig. 30 F [10]) were observed.

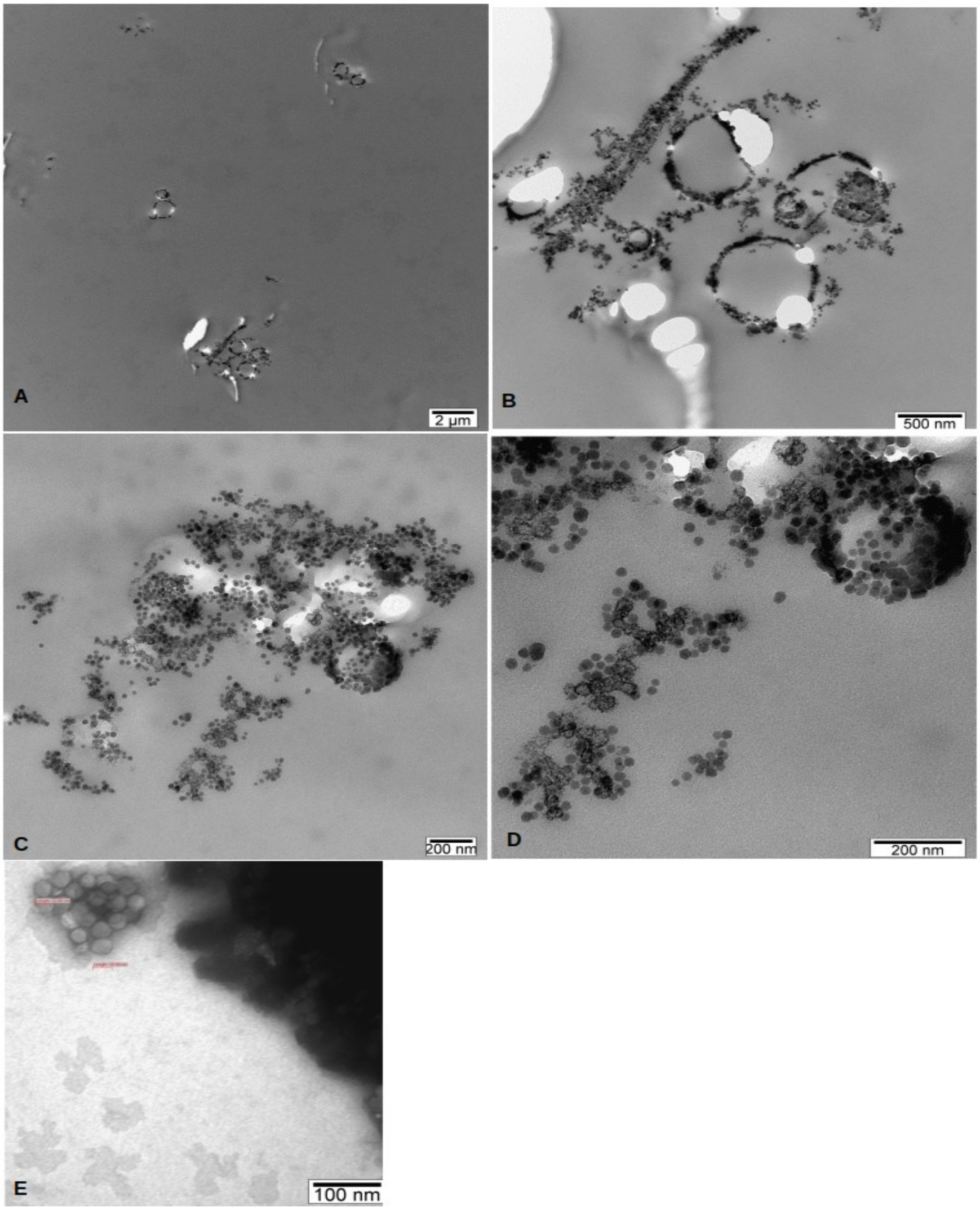


Fig. 27 1st section TEM image from absolute ethanol

TEM after flat embedding in epon counter-stained (Ur-ac and Pb)) of cell from DAL-P4 preserved in absolute ethanol at -20ºC. 1st section image with scale bar.

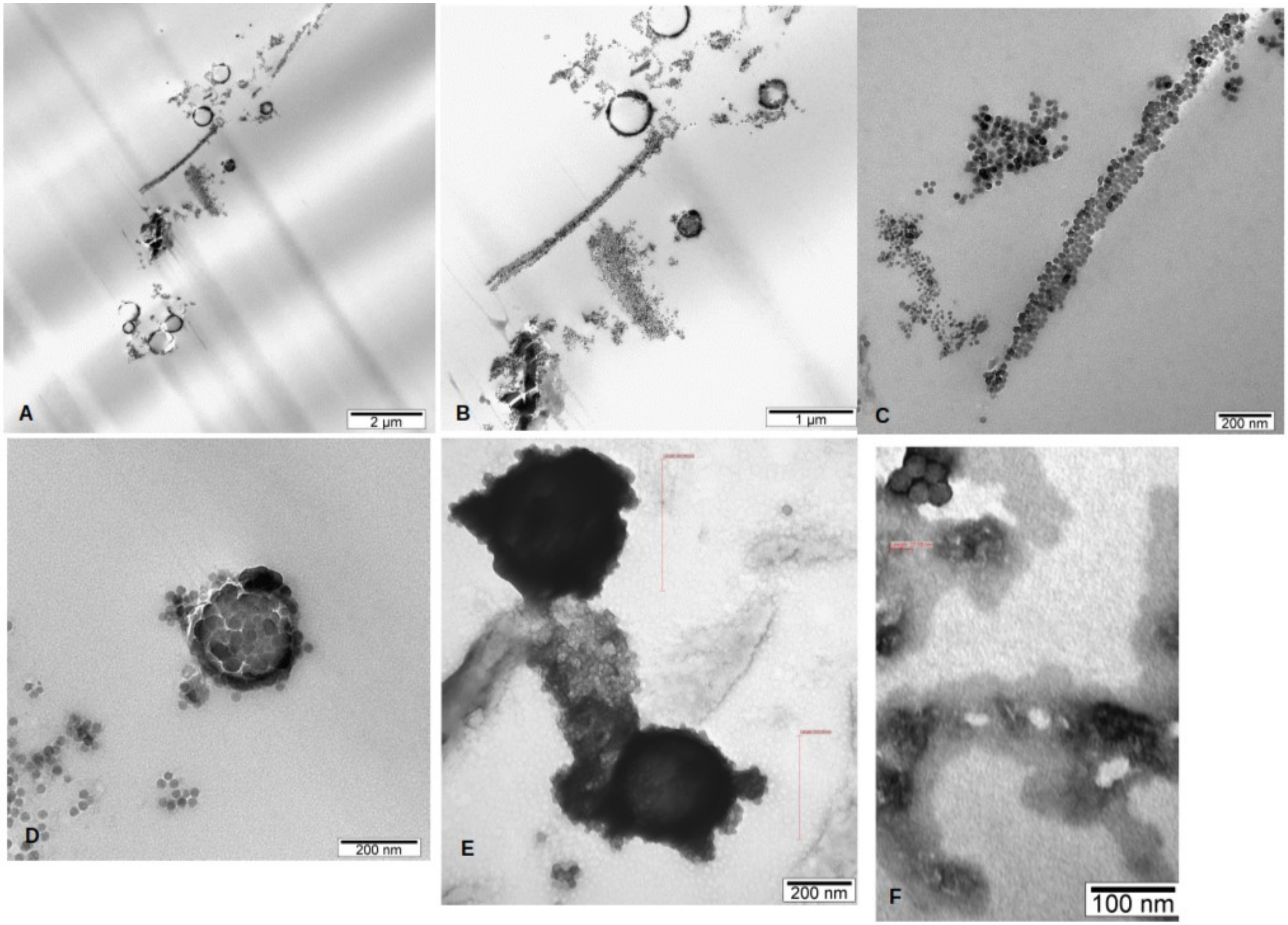


Fig. 28 1st section TEM image from Sucrose lysis buffer

1st section after flat embedding in epon counterstained (Ur-ac and Pb)) of cell from DAL-P4) preserved in Sucrose lysis buffer (pH 9) at -20°C. 1st section image with scale bar showing production of numerous EMVs (A-D). Diameter sizes of aggregates and individual EMV estimated to be between 320-410nm and 27nm respectively (E-F).

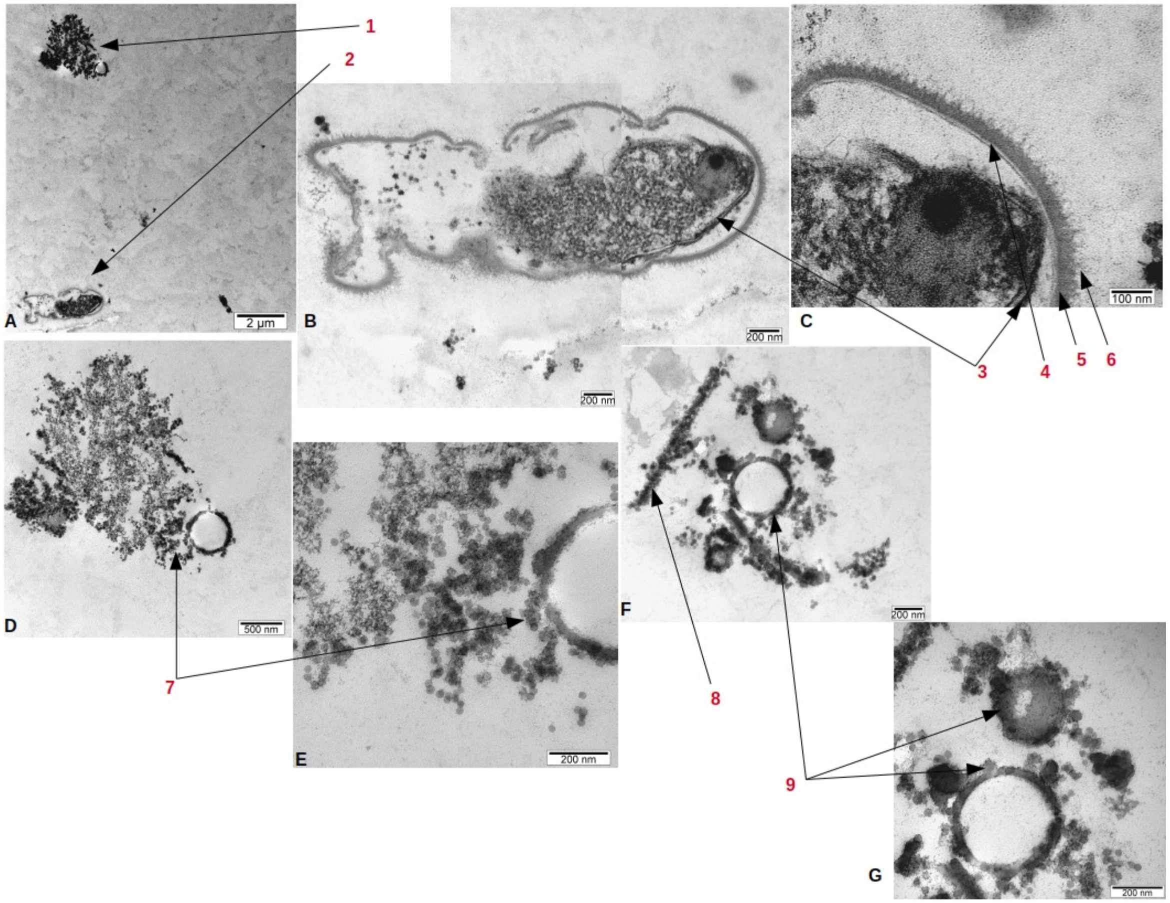


Fig. 29 2nd section TEM of cells in absolute ethanol

Images of microbial cells from DAL-P4 preserved in absolute ethanol at -20oC obtained using TEM (2nd section) after flat embedding in epon counter-stained by Ur-ac and Pb. Various coccus shaped microbe (A[1], D and E) with numerous nano-scale extrusions [7]. indeterminately shaped microbe (A[2], B and C) with shrunken plasma membrane [3], cell wall [4], capsule [5], pili [6]. Possible segemented bacteria (F[8]) and cocci bacteria (F[9] and G) with electron-dense.

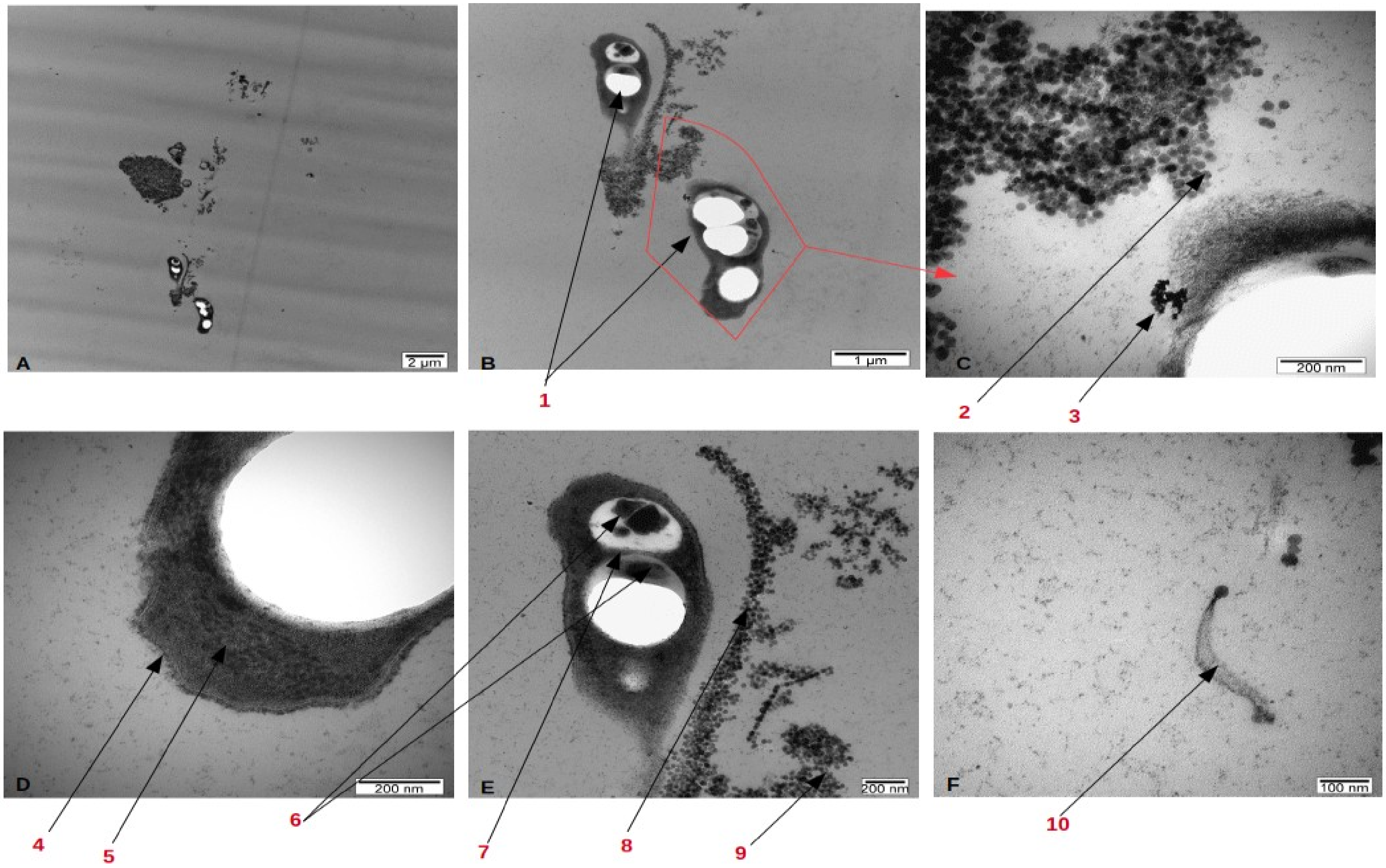


Fig. 30 2nd section TEM of cells in Sucrose lysis buffer

Images of microbial cells from DAL-P4 preserved in Sucrose lysis buffer at -20°C (A-F) obtained using TEM (2nd section after flat embedding in epon and counterstained by Ur-ac and Pb). Cocci cells undergoing cell division [1]. Magnified image showing EMVs [2] and membrane protein [3], cell membrane and polysaccharide wall [4] and ribosomes [5]. nucleoid [6], septum [7], Possible segmented bacteria with EMVs (E [8-9]). Nano fiber production by EMVs (F [10]).

5.3.2. Growth detection of microbes in special culture media and possible removal of phenolic complex

From the two special media used for growing extremophiles of DAL-P4, growth curve was plotted using the Optical Density (OD) measured over Time (T) (Appendix 12). For the acidic medium, exponential growth was observed with in 24 hrs of incubation and additional spike in OD measurement was recorded at the 96 hrs (Fig. 31). On the other hand, for the basic medium, constant increase of OD at 380 nm was observed for all the replicates of growth culture medium and the abiotic control used (Fig. 32). Production of phenol-$FeCl_3$ complex was instantaneous during the addition of $FeCl_3$ in the phenol containing special basic media and accumulation of this complex was the maximum at 72 hr after incubation (Appendix 12). The OD measurements of the replicates of inoculated basic culture media was higher than the abiotic control medium during the first 48 hrs of incubation, but started to decline after 48 hrs and reached minimum point at 120 hrs. This result can be indicative of microbial growth over the 48 hrs of incubation while significant reduction of phenol-$FeCl_3$ complex in the growth culture media was observed. The effect of conglomerations of Phenol-$FeCl_3$ complex was considered during measurement and removal of Phenol-FeCl3 complex from the basic medium was indicated after 48 hrs (Fig. 33).

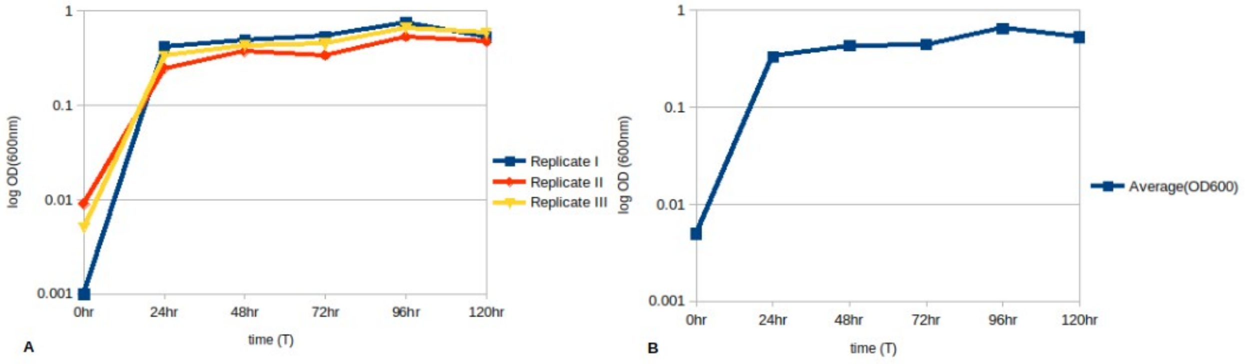


Fig. 31 Transformed log graph of $OD_{(600nm)}$ of acidic culture media

[**A**] OD of acidic culture medium with inoculum (0.1% v/v) replicates I (blue), II (red) and III (yellow) and [**B**] average OD. abiotic control was used as blank during measurement. Time (T) was measured in hour (hr) and exponential growth was reached at 48 hrs in the liquid growth medium.

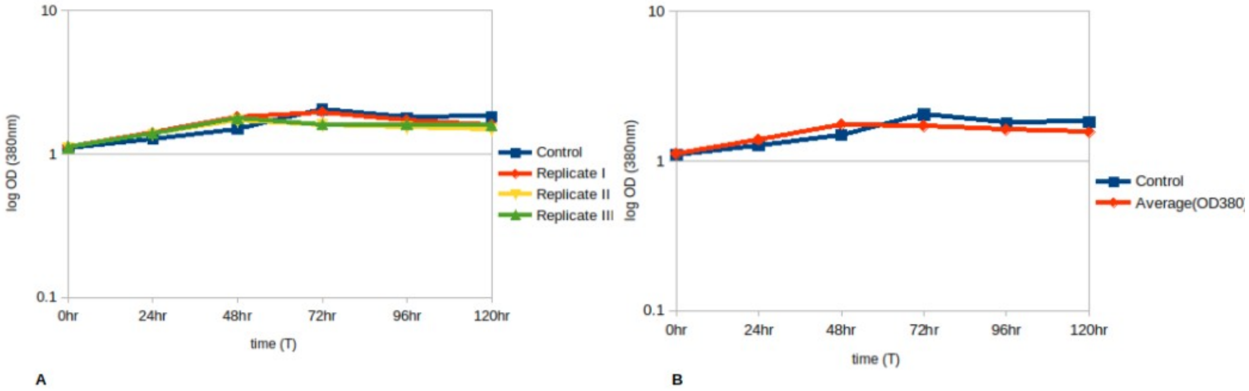


Fig. 32 Transformed log graph of $OD_{(380nm)}$ of basic culture media

[**A**] OD of basic culture medium with inoculum (0.1% v/v) replicates I (red), II (yellow) and III (green) and the abiotic control (blue). [**B**] average OD. The effect of conglomerations of Phenol-$FeCl_3$ complex was considered during measurement. Time (T) was measured in hour (hr) and growth in the 48 hrs after inoculation was indicated.

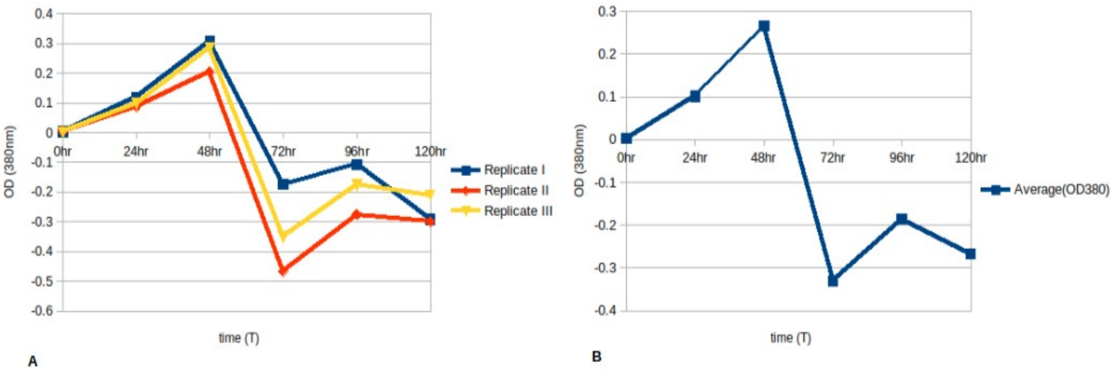


Fig. 33 OD ($_{380\ nm}$) of basic culture media with abiotic control

OD ($_{380\ nm}$) of basic culture media where abiotic control media was used as a blank. Time (T) was measured in hour (hr) and sharp decrease in measured OD was observed after 48hr of incubation at 55^{0}C and and was minimum at 72hrs. is a substrate that is used by the microbe used as usually is not the same under your sample. It can be a little thicker and thus absorb more.

Estimation of growth rate of microbes in the specialized media in terms of generation per hour was calculated using the average ODs at inoculation (T=0 hr) and at exponential growth (Fig. 31-33 and Appendix 12). The average growth rate of microbes in the acidic medium was calculated at T=0 hr (OD_0=0.005) and T=24hrs (OD_{24}=0.331) (Fig. 31 and Appendix 12). Hence, the estimated average growth rate of microbes was approximately 0.25/hour and the doubling time is approximately 4 hrs. For the basic culture, where sterile $FeCl_3$-free reference blank was used, the estimated average growth rate of microbes (exponential growth at OD_{48} = 1.756) was 0.0138/hr with 72.5 hrs as doubling time (Fig. 32 and Appendix 12). The estimated accumulation rate of oxidized phenol-$FeCl_3$ complex (OD_{48} = 1.49) was 0.009/hr. However, when the abiotic control was used as a reference blank (OD_{48} = 0.266) , the growth rate of microbes was estimated to be approximately 0.135/hr and the doubling time approximately 7.4 hrs (Fig. 33 and Appendix 12). Therefore, from these results, the growth rate of microbes in acidic media is higher than basic medium while the generation time (doubling time) was minimum.

Furthermore, during analysis of the OD/time profiles, ‘spikes’ of OD measurements were observed in Fig. 31-33. These spikes were not electronic in nature, but were “a reproducible effect" of the cells in the cuvettes. Microscopy analysis of aliquots showed high degree of cellular aggregation in both acidic and basic special culture media (Fig. 34). Viability of cells were confirmed by using Cell Tracker TM Green CMFDA fluorescent probe

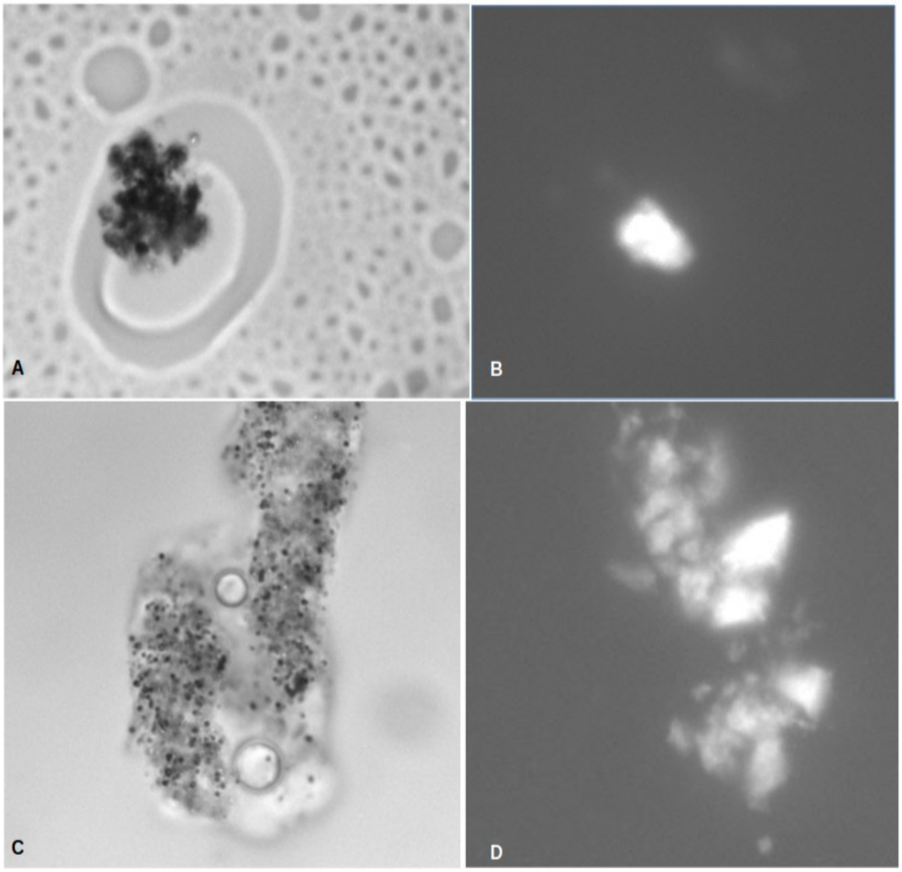


Fig. 34 None fluorescent and fluorescent cells images

None fluorescent (A and C) and fluorescent stains (B and D) of cells images (black and white). Aliquoted from acidic medium (A and B) and basic medium (C and D) were analyzed using Zeiss Axio Observer with Objective at 100x magnification under oil immersion.

5.4. Discussion

The Dallol hydrothermal field is one of the most dynamic environment on Earth (Cavalazzi *et al.,* 2019; Kotopoulou *et al.,* 2019). The total percentage of salinity recorded for DAL-P4 acid pond exceeded the 2015 measurement on the same pond by 6% just like the amounts of NO_3^- and NO_2^- which twice exceeded the previous year amount. The amount of NH_4^+ in DAL-P4 was very high, exceeding the thresholds for odor detection (approximately 1.5 mg/L), taste perception (35 mg/L) some anaerobic ground water condition (greater than 3 mg/L). Though the pH still indicated a hyper acidic condition, this type of fluctuation of physio-chemistry can cause constant stress on the inhabitant that reside in the brine of DAL.

Microorganisms are the most resilient living things on earth. Different strategies of survival have been adopted by microbial life forms in extreme environments. For instance, *Chroococcidiopsis* cyanobacteria survives the intense solar irradiation and extreme dryness of Atacama Desert by attaching to gypsum rock to extract water and create a thin layer of water-containing calcium sulfate-based mineral (Huang *et al.*, 2020). On the other hand, Psychrophiles survive very cold temperature by adjusting the surface area of their proteins to be more flexible (Shokrollahzade *et al.*, 2015). In addition, some representatives of the Phylum Firmicutes produces endospores and enter to metabolically inactive state to survive single or multiple limiting environmental conditions (Filippidou *et al.*, 2016). In the case of extremophiles of the studied Dallol acid pools, we've found microscopic evidence that at least some extremophiles produce EMV.

The centrifugation method was reliable for collecting cells from brine samples for direct visualization of the cells as reported in Silva *et al.*, (2013). Based on our TEM analysis, majority of the EMVs observed had a single-bilayer structure with diameter less than 30 nm (nano-scaled) and were originated from the plasma membrane. Important inter-cellular communications between cells are known to be served by EMVs including transfer of proteins, lipids, and nucleic acids (Raposo and Stoorvogel, 2013; Go *et al.*, 2019; Jahromi and Fuhrmann, 2021). Evidences of protective roles of eukaryotic or prokaryotic EMVs against both internal and external stressors are well documented (Turturici *et al.,* 2014; Yokoyama *et al.,* 2017; Hosseinkhani *et al.*, 2017; Go *et al.*, 2019; Baeza and Mercade, 2021). In some cold-adapted Antarctic Gram-negative, high levels of extracellular ATP were found to be secreted in EMVs, which are associated with significant increase in the speed and amount of biofilm formation (Baeza and Mercade, 2021). Although there is a gap of information about the cargo packaged within the EMV of Dallol sample, their physiological relevance and involvement can be predicted as stress relief mechanism, especially to the lysing effects of different chemicals used for DNA extraction (Tilahun *et al.,* 2021[a]).

Aggregates of nano-scaled EMV covering cells were clearly observed in our TEM analysis. Different sectioning of TEM revealed cellular characteristics that assist surviving in the preserving and lysing conditions of buffers that are known for their detrimental effects on cells (Mitchell and Takacs-Vesbach, 2008). Images of intact and sometimes actively dividing cells protected by the EMVs were revealed inside the lysing buffer. The potential use of nano scaled EMVs for drug delevery in the field of Medicine has been well documented in Herrmann *et al.,* (2021). Other important cellular features of the extremophiles such as the cell capsule and membrane as organelles such as ribosomes were clearly identified by the TEM.

Further investigation to answer the question of viability of cells inside the cell lysing buffer also revealed the capability of the extremophiles of DAL-P4 sample surviving and adapting the extreme conditions of the special growth medium. Even though DAL-P4 is naturally very acidic, growth was observed in the alkaline liquid medium used in this investigation. Yet, growth rate of microbes in the acidic special culture medium was much higher than the basic culture media. Tolerance of wide pH range by extremophiles under laboratory conditions have been broadly reported. Maintaining the cytoplasmic and vacuolar pH constant is important for safekeeping intracellular metabolism (Slonczewski *et al.,* 1987; Hesse *et al.,* 2002; Dhakar *et al.,* 2014). This type of cellular adaptation is usually accompanied by stimulation of specific operons, transcriptional regulations and synthesis of proteins such as 36 kDa protein (p36) (Amaro *et al*., 1991; Watson *et al.,* 1992; Hughes *et al.,* 2007; Dhakar *et al.,* 2014).

The high-resolution microscopy and staining efficiency of Cell Tracker TM Green CMFDA (or 5-chloromethylfluorescein diacetate) confirmed the presence of viable cells as well as biofilm in the two types special media prepared. The CMFDA used is an artificial substrate for esterase and used to stain bacterial communities where intracellular esterase activity can be detected (Fuller *et al.,* 2000; Senjarini *et al.,* 2013). Principally, CMFDA bound to the glutathion S-transferase (GST) to the cell's protein pools (Senjarini *et al.,* 2013). GST is known to be highly active in natural biofilms from extreme environments or at stressful laboratory conditions (Hoque *et al.,* 2007; Al-Madboly *et al.,* 2020). Though the labeling capacity of CMFDA is shorter (maximum of 2-3 days), it effectively stained the cells' "biofilm" without compromising cell viability or altering cell adhesion characteristics (Fuller *et al.,* 2000).

From biotechnological point of view, the evidence of production of numerous EMV by DAL-P4 extremophiles was quite compelling. Especially their stability under the harsh environmental conditions such as the acid pool of Dallol can have great prospect in industrial application such as in Medical, Food, and other fields (Wiklander *et al.,* 2019; de Paula *et al.,* 2019; Morales *et al.,* 2021). The removal of phenolic complex observed during growth of microbes under basic condition is also an indication of the potential of these extremophiles in bioremediation processes.

5.5. Conclusion

The resulted TEM showed production of nano scaled EMVs, which have been associated with cell’s response to external stress and other physiological activities. From the obtained image, the production of EMVs in Dallol sample is at least one factor that protected the cells from disintegrating by the action of the lysis buffer and continued to be viable and proliferating. Further study on these extremophiles will create better understanding to design and develop steps required to utilize their full potential in medical and environmental biotechnology.

Chapter Six

6. General Conclusion and Recommendation

The current studies described here contribute to our understanding the biodiversity of poly-extreme aquatic environments in Danakil Depression by expanding our knowledge on the prokaryotic phylogenetic diversity that exists in the ponds of DAL, GAL, BLA, MUP as well as selected coastal sites of Lake As’ale. Further assessment on metabolic pathways also revealed the potential nutrient cycling pathways and metabolic networks of omnipresent, uncultured microbial populations in the studied sites.

Studying the microbial diversities in DAL, BLA, GAL, MUP and the three sampling sites of coastal parts of Lake As’ale using whole metagenome sequencing approach, helped to determine prokaryotic community composition of the poly-extreme environments of Danakil Depression. Moreover, in chapter 3, we reported hydrochemistry of brine samples collected in 2015 from DAL, BLA, GAL, MUP and three sites of coastal parts of Lake As’ale. Based on the results, DAL, BLA and GAL sample sites were classified collectively as Hypersaline Acid Ponds (HsAP), while MUP and Lake As’ale were classified together as Hypersaline Aquatic Systems (HsAqS). Furthermore, all the studied sites except for DAL were found to be fed by non-volcanic, dominantly rainfall originated groundwater, while DAL had strong volcanic signature. We also demonstrated the existence of methodological and systematic bias in the extraction of environmental DNA from extremophiles by use of inadequate chemical extraction methods. We identified that the CTAB-SDS DNA extraction method that would accurately lyse most mesophilic microbial cells to recover the genomic materials was inadequate for the extremophiles from HsAP of this study. Instead, the combination of chemical and physical (specifically lysing of cells using sonication) extraction methods provided sufficient DNA material for the metagenome study.

Therefore, multifaceted DNA extraction methods, NGS and a non-assembled metagenome analysis tool (MetaPhlAn 2) were used to survey the prokaryotic community composition of the poly-extreme environments of Danakil Depression. As a result, diverse and ubiquitous distribution of bacterial and archaeal phyla were documented in the HsAp and HsAqS, respectively. The described results provide taxonomy profile for each sample sites at species level and many of the identified OTUs were unclassified at different taxonomic level. Nevertheless, the taxonomic families Bradyrhizobiaceae and Burkholderiaceae were dominant in HsAP, while Halobacteriaceae and Haloferacaceae were dominant in Lake As’ale and MUP.

The annotation of important gene from metagenome sequence data and mapping in nutrient metabolism pathways, helped to predict nutrient cycling in in Danakil Depression. In Chapter 4, we described the predicted metabolic pathways of the sample sites using in silico analysis in consideration to nutrients cycling. The investigation, reconstruction and visualization of metabolic networks of carbohydrate, nitrogen, sulfur and phosphorous in the studied sites allowed us to easily comprehend the model and perform its graphical analysis. The different predicted homologous protein families involved in the four nutrient bio-cycling helped us to identify the parts of the metabolic systems with a designated function. The sources of predicted genes, especially mapped to N_2 and CO_2 cycles, were various individual OTUs but belonging to the dominant taxonomic families profiled indicating cooperative (syntrophic) interactions among inhabiting microbes. The genomic data that revealed the syntrophic interactions and contribution of individual OTUs in the N_2 and CO_2 fixation pathways can be used as basic information to design applications in industrial, agricultural as well as environmental biotechnology under hypersaline and acidic conditions. In addition, the results in Chapter 3 and 4 give insights for future studies that target microbial communities and/or genes with possible potential in the fields of synthetic biology as well as bio-fuel production.

Further investigation on extremophiles of DAL sample showed production of EMVs as stress relief mechanism and the removal of phenolic complex under stressful conditions. These data showed the great biotechnological potential of extremophiles of DAL in the fields of medicine and environment among others.

The overall investigations conducted clearly showed that the acid ponds and the hypersaline aquatic systems scattered in the Danakil Depression are ideal places to study how the poly-extreme environments affect organic chemistry and life on Earth, or on other sites in the solar system. Using molecular methods, novel populations of different Bacteria and Archaea occupying the poly-extreme environments were identified. Moreover, the presence of possible interaction of microbial communities within their respective habitats helped to understand how individual microbial group contribute to maintain ecosystem fitness and resist environmental stress.

In general, the investigations conducted showed that the extremophilic organisms of the studied poly extreme environments of Danakil Depression have huge potential in different biotechnological fields such as agriculture, environmental, medical as well as different industries for production of consumable products. Hence further investigations should be conducted on specific enzymes and other bio-catalysts identified from the in silico results, using cloning and other life technologies to design, synthesize, test and get better variants with high expression and capacity without isolation of extremophiles.

7. References

1. Abbate E., Bruni P., and Sagri M. (2015). Geology of Ethiopia: A Review and Geomorphological Perspectives. In Landscapes and Landforms of Ethiopia pp.33-64. *Springer*. The Netherlands. DOI:10.1007/978-94-017-8026-1_2
2. Abram K., Udaondo Z., Bleker C., Wanchai V., Wassenaar T. M., Robeson M. S. and Ussery D. W. (2020).What can we learn from over 100,000 *Escherichia coli* genomes?. *BioRxiv,* **708131,** DOI:10.1101/708131
3. Abreu F., Leao P., Vargas G., Cypriano J., Figueiredo V., Enrich-Prast A., Bazylinski D.A. and Lins U. (2018). Culture-independent characterization of a novel magnetotactic member affiliated to the Beta class of the Proteobacteria phylum from an acidic lagoon. *Environmental Microbiology,* **20(7),** 2615-2624. DOI:10.1111/1462-2920.14286
4. Ackermann H. and Kropinski A.M. (2007). Curated list of prokaryote viruses with fully sequenced genomes. *Research in Microbiology,* **158(7),** 555-566. DOI:10.1016/j.resmic.2007.07.006
5. Adamczack J., Hoffmann M., Papke U., Haufschildt K., Nicke T., Bröring M., Sezer M., Weimar R., Kuhlmann U., Hildebrandt P., and Layer G. (2014). NirN from *Pseudomonas aeruginosa* is a novel electron-bifurcating dehydrogenase catalyzing the last step of heme d1 biosynthesis. *Journal of Biological Chemistry,* **289(44),** 30753-30762. DOI: 10.1074/jbc.M114.603886
6. Ahrens C. D. (2011). Essentials of meteorology: an invitation to the atmosphere. *6th ed*. pp. 60. Brooks/Cole: Belmont, CA, USA.
7. Akhtar M., Ahmad N. and Booij M.J. (2008). The impact of climate change on the water resources of Hindukush–Karakorum–Himalaya region under different glacier

coverage scenarios. *Journal of Hydrology,* **355 (1–4),** 148-163. DOI:10.1016/j.jhydrol.2008.03.015

8. Al-Madboly L. A., Ali S. M., Fakharany E., Ragab A. E., Khedr E. G., and Elokely K. M. (2020). Stress-Based Production, and Characterization of Glutathione Peroxidase and Glutathione S-Transferase Enzymes From *Lactobacillus plantarum. Frontiers in Bioengineering and Biotechnology,* **8,**78. DOI:10.3389/fbioe.2020.00078
9. Alkan C., Sajjadian S., and Eichler E. E. (2011). Limitations of next-generation genome sequence assembly. *Nature methods,* **8(1),** 61–65. DOI:10.1038/nmeth.1527
10. Allaband C., McDonald D., Vázquez-Baeza Y., Minich J. J., Tripathi A., Brenner D. A., Loomba R., Smarr L., Sandborn W. J., Schnabl B., Dorrestein P., Zarrinpar A., and Knight R. (2019). Microbiome 101: Studying, Analyzing, and Interpreting Gut Microbiome Data for Clinicians. *Clinical Gastroenterology and Hepatology,* **17(2),** 218–230. DOI:10.1016/j.cgh.2018.09.017.
11. Alneberg J., Bjarnason B.S., Bruijn I., Schirmer M., Quick J., Ijaz U.Z., Lahti L., Loman N.J., Andersson A.F. and Quince C. (2014). Binning metagenomic contigs by coverage and composition. *Nature Methods,* **11(11),**1144–1146. DOI:10.1038/nmeth.3103
12. Alves Ld, F., Silva-Rocha, R., and Guazzaroni, M.E. (2017). Enhancing Metagenomic Approaches Through Synthetic Biology. In: Charles T., Liles M., Sessitsch A. (eds) Functional Metagenomics: Tools and Applications. Springer, Cham. DOI: 10.1007/978-3-319-61510-3_5
13. Amaro A M , Chamorro D , Seeger M, Arredondo R, Peirano I. and Jerez C. A. (1991). Effect of external pH perturbations on *in vivo* protein synthesis by the

acidophilic bacterium *Thiobacillus ferrooxidans. Journal of Bacteriology,* **173(2),** DOI: 10.1128/jb.173.2.910-915.1991

14. Antón J., Oren A., Benlloch S., Rodríguez-Valera F., Amann R. and Rosselló-Mora R. (2002). *Salinibacter rube*r gen. nov., sp. nov., a novel, extremely halophilic member of the Bacteria from saltern crystallizer ponds. *International journal of systematic and evolutionary microbiology,* **52(2),** 485-491. DOI:10.1099/00207713-52-2-485
15. Asnicar F., Weingart G., Tickle T. L., Huttenhower C., and Segata, N. (2015). Compact graphical representation of phylogenetic data and metadata with GraPhlAn, *PeerJ.* **3,** e1029. DOI:10.7717/peerj.1029
16. Asrat A. (2016). The Danakil Depression: An exceptional place where different types of extreme environments coexist. In Africa’s Top Geological Sites. Pp 189-196. Struik Nature, South Africa.
17. Baas-Becking L. G. M. (1931). Salt effects on swarmers of *Dunaliella viridisteod. The Journal of General Physiology,* **14 (6),** 765. DOI: 10.1085/jgp.14.6.765
18. Badhai J., Ghosh T. S., and Das, S. K. (2015). Taxonomic and functional characteristics of microbial communities and their correlation with physicochemical properties of four geothermal springs in Odisha, India. *Frontiers in microbiology,* **6,** 1166. DOI:10.3389/fmicb.2015.01166
19. Baeza N. and Mercade E. (2021). Relationship Between Membrane Vesicles, Extracellular ATP and Biofilm Formation in Antarctic Gram-Negative Bacteria. *Microbial Ecology,* **81(3),**1-12. DOI: 10.1007/s00248-020-01614-6
20. Bag S., Saha B., Mehta O, Anbumani D., Kumar N., Dayal M., Pant A., Kumar P., Saxena S., Allin K.H., Hansen T., Arumugam M., Vestergaard H., Pedersen O.,

Pereira V., Abraham P., Tripathi R., Wadhwa N., Bhatnagar S., Prakash V.G., Radha V., Anjana R.M., Mohan V., Takeda K., Kurakawa T., Nair G.B. and Das B. (2016). An Improved Method for High Quality Metagenomics DNA Extraction from Human and Environmental Samples. *Scientific reports,* **31(6),** 26775. DOI: 10.1038/srep26775

21. Bağcı C., Beier S., Górska A., Huson D.H. (2019). Introduction to the Analysis of Environmental Sequences: Metagenomics with MEGAN. In: Anisimova M. (eds) Evolutionary Genomics. Methods in Molecular Biology, vol 1910. Humana, New York, NY. DOI: 10.1007/978-1-4939-9074-0_19.
22. Baker, M. (2012). Structural variation: the genome's hidden architecture. *Nature methods,* **9(2),** 133–137. DOI:10.1038/nmeth.1858
23. Baker-Austin C. and Dopson M. (2007). Life in acid: pH homeostasis in acidophiles. *Trends in microbiology*, **15(4),** 165-171. DOI: 10.1016/j.tim.2007.02.005
24. Bala A. and Singh B. (2019). Cellulolytic and xylanolytic enzymes of thermophiles for the production of renewable biofuels. *Renewable Energy,* **136,** 1231-1244. DOI:10.1016/j.renene.2018.09.100
25. Banda J. F., Lu Y., Hao C., Pei L., Du Z., Zhang Y., Wei P. and Dong H. (2020). The Effects of Salinity and pH on Microbial Community Diversity and Distribution Pattern in the Brines of Soda Lakes in Badain Jaran Desert, China. *Geomicrobiology Journal,* **37(1),** 1-12. DOI: 10.1080/01490451.2019.1654568
26. Bang C., Dagan T, Deines P., Dubilier N., Duschl W. J., Fraune S., Hentschel U., Hirt H., Hülter N., Lachnit T., Picazo D., Pita L., Pogoreutz C., Rädecker N., Saad M. M., Schmitz R. A., Schulenburg H., Voolstra C. R., Weiland-Bräuer N., Ziegler M. and Bosch T. C.G. (2018). Metaorganisms in extreme environments: do microbes

play a role in organismal adaptation? *Zoology,* **127,** 1-19. DOI:10.1016/j.zool.2018.02.004

27. Barba M., Czosnek H., and Hadidi A. (2014). Historical perspective, development and applications of next-generation sequencing in plant virology. *Viruses,* **6(1),** 106–136. DOI:10.3390/v6010106
28. Barbieri R., and Cavalazzi B. (2014). How Do Modern Extreme Hydrothermal Environments Inform the Identification of Martian Habitability? The Case of the El Tatio Geyser Field. *Challenges*. **5(2)**: 430-443. DOI:10.3390/challe5020430
29. Barnard A.S. (2014). In Silico Veritas. *ACS Nano,* **8(7),** 6520–6525. DOI: 10.1021/nn502808y
30. Barriuso J. and Martínez M. J.(2015). In Silico Metagenomes Mining to Discover Novel Esterases with Industrial Application by Sequential Search Strategies. *Journal of Microbiology and Biotechnology,* **25(5),** 732-737. DOI:10.4014/jmb.1406.06049
31. Bassel G.W., Holdsworth M.J. and Provart N.J. (2011). Seed Bioinformatics. *Methods in Molecular Biology*, **773,** 403-19. DOI: 10.1007/978-1-61779-231-1_23
32. Bedoya K., Coltell O., Cabarcas F. and Alzate J. F. (2019). Metagenomic assessment of the microbial community and methanogenic pathways in biosolids from a municipal wastewater treatment plant in Medellín, Colombia. S*cience of The Total Environment*, **648,** 572-581. DOI:10.1016/j.scitotenv.2018.08.119
33. Behjati S. and Tarpey P. S. (2013). What is next generation sequencing? *Archives of Disease in Childhood. Education and Practice Edition,* **98(6),** 236–238. DOI:10.1136/archdischild-2013-304340

34. Belilla J., Moreira D., Jardillier L., Reboul G., Benzerara K, López-García J.M., Bertolino P., López-Archilla A.I. and López-García P. (2019). Hyperdiverse archaea near life limits at the polyextreme geothermal Dallol area. *Nature Ecology and Evolution,* **3,** 1552–1561. DOI:10.1038/s41559-019-1005-0

35. Berg G., Kusstatscher P., Abdelfattah A., Cernava T. and Smalla K. (2021). Microbiome Modulation—Toward a Better Understanding of Plant Microbiome Response to Microbial Inoculants. *Frontiers in Microbiology,* **12,** 650610. DOI:10.3389/fmicb.2021.650610

36. Berg G., Rybakova D., Fischer D., Cernava T., Vergès M.C., Charles T., Chen X., Cocolin L, Eversole K., Corral G.H., Kazou M., Kinkel L., Lange L., Lima N., Loy A., Macklin J.A., Maguin E., Mauchline T., McClure R., Mitter B., Ryan M., Sarand I., Smidt H., Schelkle B., Roume H., Kiran G.S., Selvin J., de Souza R.S.C., Overbeek L., Singh B.K., Wagner M., Walsh A., Sessitsch A. and Schloter M. (2020). Microbiome definition re-visited: old concepts and new challenges. *Microbiome,* **8(1),** 103. DOI:10.1186/s40168-020-00875-0

37. Berghuis B.A, Yu F.B., Schulz F., Blainey P.C., Woyke T. and Quake S.R. (2019). *Hydrogenotrophic methanogenesis* in archaeal phylum Verstraetearchaeota reveals the shared ancestry of all methanogens. *Proceedings of the National Academy of Sciences,* **116(11),** 5037-5044. DOI: 10.1073/pnas.1815631116

38. Beyth M. (1978). A comparative study of the sedimentary fills of the Danakil Depression (Ethiopia) and Dead-Sea Rift (Israel). *Tectonophysics,* **46(3-4)**, 357-367. DOI:10.1016/0040-1951(78)90213-5

39. Bhatia R.K., Ullah S., Hoque M.Z., Ahmad I., Yang Y., Bhatt A.K. and Bhatia S.K. (2021). Psychrophiles: A source of cold-adapted enzymes for energy efficient

biotechnological industrial processes. *Journal of Environmental Chemical Engineering,* **9(1),** 104607. DOI: 10.1016/j.jece.2020.104607

40. Black K. A. and Dos Santos P. C. (2015). Shared-intermediates in the biosynthesis of thio-cofactors: mechanism and functions of cysteine desulfurases and sulfur acceptors. *Biochimica et Biophysica Acta,* **1853(6),** 1470–1480. DOI:10.1016/j.bbamcr.2014.10.018
41. Bolger A. M., Lohse M., and Usadel B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics,* **30(15),** 2114–2120. DOI: 10.1093/bioinformatics/btu170
42. Boor K. J. (2006). Bacterial Stress Responses: What Doesn't Kill Them Can Make Them Stronger. *PLoS Biology,* **4(1),** e23. DOI:10.1371/journal.pbio.0040023
43. Bragg L. and Tyson G.W. (2014). Metagenomics using next-generation sequencing. *Methods in molecular biology,* **1096,** 183-201. DOI: 10.1007/978-1-62703-712-9_15
44. Bragina A., Oberauner-Wappis L., Zachow C., Halwachs B., Thallinger G.G., Müller H. and Berg, G. (2014). The Sphagnum microbiome supports bog ecosystem functioning under extreme conditions. *Molecular ecology,* **23(18),** 4498-4510. DOI:10.1111/mec.12885
45. Braüer S.L., Cadillo-Quiroz H., Yashiro E., Yavitt and Zinder S.H. (2006). Isolation of a novel acidiphilic methanogen from an acidic peat bog. *Nature,* **442,** 13. DOI:10.1038/nature04810
46. Bray, J.R. and Curtis, J.T. (1957). An Ordination of the Upland Forest Communities of Southern Wisconsin. *Ecological Monographs,* **27,** 325-349. DOI:10.2307/1942268

47. Brüggemann, H., and Lood, R. (2013). Bacteriophages infecting *Propionibacterium acnes. BioMed Research International*, **2013,** 705741. DOI:10.1155/2013/705741

48. Buchfink B., Xie C. and Huson D. H. (2015). Fast and sensitive protein alignment using DIAMOND. *Nature Methods,* **12,** 59–60. DOI: 10.1038/nmeth.3176

49. Bunge J., Willis A. and Walsh F. (2014). Estimating the number of species in microbial diversity studies. *Annual Review of Statistics and Its Application,* **1,** 427–445. DOI:10.1146/annurev-statistics-022513-115654

50. Burén S., López-Torrejón G. and Rubio L. M. (2018). Extreme bioengineering to meet the nitrogen challenge. *Proceedings of the National Academy of Sciences,* **115 (36),** 8849-8851. DOI: 10.1073/pnas.1812247115

51. Burns D., Camakaris H., Janssen P., and Dyall-Smith M. (2004). Combined use of cultivation-dependent and cultivation-independent methods indicates that members of most haloarchaeal groups in an Australian crystallizer pond are cultivable. *Applied and Environmental Microbiology,* **70(9),** 5258-5265. DOI: 10.1128/AEM.70.9.5258-5265.2004

52. Byvalov A. A., Konyshev I.V., Uversky V.N., Dentovskaya S.V., Anisimov A.P. (2020). *Yersinia* outer membrane vesicles as potential vaccine candidates in protecting against plague. *Biomolecules,* **10(12),** 1694. DOI:10.3390/biom10121694

53. Cabrera M.Á. and Blamey J.M. (2018). Biotechnological applications of archaeal enzymes from extreme environments. *Biological Research,* **51(1),** 37. DOI:10.1186/s40659-018-0186-3

54. Cáceres-Moreno P., Muñoz-Ibacache S. A., Monsalves M. T., Amenabar M. J. and Blamey J. M. (2019). Functional Approach for the Development and Production of

Novel Extreme Biocatalysts. *Next Generation Biomanufacturing Technologies*, **1329,** 1-22. DOI:10.1021/bk-2019-1329.ch001

55. Carr S., Schubotz F., Dunbar R., Mills C. T., Dias R., Summons R. E. and Mandernack K. W. (2018). Acetoclastic *Methanosaeta* are dominant methanogens in organic-rich Antarctic marine sediments. *The ISME Journal,* **12(2),** 330–342. DOI:10.1038/ismej.2017.150
56. Carrier T. J. and Reitzel A. M. (2017). The hologenome across environments and the implications of a host-associated microbial repertoire. *Frontiers in Microbiology,* **8,** 882. DOI:10.3389/fmicb.2017.00802
57. Carrizo D., Sánchez-García L., Rodriguez N. and Gómez F. (2019). Lipid biomarker and carbon stable isotope survey on the Dallol hydrothermal system in Ethiopia. *Astrobiology,* **19(12),** 1474-1489. DOI:10.1089/ast.2018.1963
58. Cavalazzi B., Barbieri R., Gómez F., Capaccioni B., Olsson-Francis K., Pondrelli M., Rossi A. P., Hickman-Lewis K., Agangi A., Gasparotto G., Glamoclija M., Ori G. G., Rodriguez N. and Hagos M. (2019).The Dallol geothermal area, northern Afar (Ethiopia): An exceptional planetary field analog on Earth. *Astrobiology,* **19(4),** 553–578. DOI:10.1089/ast.2018.1926
59. Cavicchioli R., Siddiqui K. S., Andrews D. and Sowers K. R. (2002). Low-temperature extremophiles and their applications. *Current opinion in biotechnology,* **13(3),** 253-61. DOI:10.1016/s0958-1669(02)00317-8
60. Cecchini G., Schröder I., Gunsalus R. P. and Maklashina E. (2002). Succinate dehydrogenase and fumarate reductase from *Escherichia coli*. *Biochimica et biophysica acta,* **1553(1-2),** 140-157. DOI:10.1016/s0005-2728(01)00238-9

61. Chakravorty D., Khan M. F. and Patra, S. (2017). Multifactorial level of extremostability of proteins: can they be exploited for protein engineering?. *Extremophiles*, **21(3),** 419–444. DOI:10.1007/s00792-016-0908-9
62. Chapman J., Ismail A. E and Dinu C. Z. (2018). Industrial applications of enzymes: recent advances, techniques, and outlooks. *Catalysts,* **8(6),** 238. DOI:10.3390/catal8060238
63. Chen Z., Zheng Y., Ding C., Ren X., Yuan J., Sun F. and Li Y. (2017). Integrated metagenomics and molecular ecological network analysis of bacterial community composition during the phytoremediation of cadmium-contaminated soils by bioenergy crops. *Ecotoxicology and Environmental Safety,* **145,** 111-118. DOI:10.1016/j.ecoenv.2017.07.019
64. Chen M. Y., Tsay S. S., Chen K. Y., Shi Y. C., Lin Y. T. and Lin G. H. (2002). *Pseudoxanthomonas taiwanensis sp. nov.*, a novel thermophilic, N_2O-producing species isolated from hot springs. *International Journal of Systematic and Evolutionary Microbiology,* **52(6),** 2155-61. DOI:10.1099/00207713-52-6-2155
65. Cheng M., Cao L. and Ning K. (2019). Microbiome big data mining and applications using single cell technologies and metagenomics approaches toward precision medicine. *Frontiers in Genetics,* **10,** 972. DOI:10.3389/fgene.2019.00972
66. Chitra K. and Maheswari D. (2017). A comparative study of various clustering algorithms in data mining. *International Journal of Computer Science and Mobile Computing,* **6(8),** 109-115. ISSN 2320–088X
67. Coker J. A. (2016). Extremophiles and biotechnology: current uses and prospects. *F1000Research,* ***5,*** F1000Res Faculty Rev-396. DOI:10.12688/f1000research.7432.1

68. Compeau P., Pevzner P. and Tesler G. (2011). How to apply de Bruijn graphs to genome assembly. *Nature Biotechnology,* **29,** 987-91. DOI:10.1038/nbt.2023

69. Cowan D. A., Ramond J-B., Makhalanyane T. P. and de Maayer P. (2015). Metagenomics of extreme environments. *Current Opinion in Microbiology,* **25,** 97-102. DOI: 10.1016/j.mib.2015.05.005

70. Crognale S., Venturi S., Tassi F., Rossetti S., Rashed H., Cabassi J., Capecchiacci F., Nisi B., Vaselli O., Morrison H. G., Sogin M. L. and Fazi S. (2018). Microbiome profiling in extremely acidic soils affected by hydrothermal fluids: the case of the solfatara crater (Campi Flegrei, southern Italy). *FEMS Microbiology Ecology,* **94,**12. DOI:10.1093/femsec/fiy190

71. Cui H., Lin Z., Dong Y., Zhou P. and Liu S. (2007). *Halorubrum litoreum* sp. nov., an extremely halophilic archaeon from a solar saltern. *International Journal of Systematic and Evolutionary Microbiology,* **57(10),** 2204–2206. DOI:10.1099/ijs.0.65268-0

72. Cummins P. L., Kannappan B. and Gready J.bE. (2018). Directions for optimization of photosynthetic carbon fixation: RuBisCO's efficiency may not be so constrained after all. *Frontiers in Plant Science,* **9,** 183. DOI:10.3389/fpls.2018.00183

73. Dana H., Mahmoodi Chalbatani G., Gharagouzloo E., Miri S. R., Memari F., Rasoolzadeh R., Zinatizadeh M. R., Kheirandish Z. P. and Marmari V. (2020). In silico analysis, molecular docking, molecular dynamic, cloning, expression and purification of chimeric protein in colorectal cancer treatment. *Drug Design, Development and Therapy,* **14**, 309–329. DOI:10.2147/DDDT.S231958

74. Darling A. E., Jospin G., Lowe E., Matsen F. A., Bik H. M. and Eisen J. A. (2014). PhyloSift: phylogenetic analysis of genomes and metagenomes. *PeerJ*, **2,** e243. DOI:10.7717/peerj.243

75. Davies K. (2017). Ring in the New: Rady children's hospital introduces NovaSeq for newborn diagnoses. *Clinical OMICs,* **4(6),** 16-17. DOI:10.1089/clinomi.04.06.12

76. Deamer D. and Damer B. (2017). Can Life Begin on Enceladus? A perspective from hydrothermal chemistry. *Astrobiology,* **17(9),** 834-839. DOI:10.1089/ast.2016.1610

77. de Bakker J., LaMarre J., Papangelakis V. and Davis B. (2011). HCl leaching and acid regeneration using $MgCl_2$ brines and molten salt hydrates. *TMS Annual Meeting,* **1,** 521-528. DOI:10.1002/9781118495285.ch59

78. de Paula R. G., Antoniêto A. C. C., Nogueira K. M. V., Ribeiro L. F. C., Rocha M. C., Malavazi I., Almeida F. and Silva R. N. (2019). Extracellular vesicles carry cellulases in the industrial fungus *Trichoderma reesei. Biotechnology for Biofuels,* **12,** 146. DOI:10.1186/s13068-019-1487-7

79. Dhakar, K., Sharma, A. and Pandey, A. (2014). Cold, pH and salt tolerant *Penicillium* spp. inhabit the high altitude soils in Himalaya, India. *World Journal of Microbiology and Biotechnology,* **30(4),** 1315–1324. DOI:10.1007/s11274-013-1545-4

80. Dixon R., Cheng Q., Shen G., Day A. and Dowson-Day M. (1997). Nif gene transfer and expression in chloroplasts: Prospects and problems. *Plant and Soil,* **194,** 193-203. DOI:10.1023/A:1004296703638

81. Donato P. D., Buono A., Poli A., Finore I., Abbamondi G. R., Nicolaus B. and Lam L. (2019). Exploring marine environments for the identification of extremophiles

and their enzymes for sustainable and green bioprocesses. *Sustainability,* **11,** 149. DOI:10.3390/su11010149

82. Dou Z., Heinhorst S., Williams E. B., Murin C. D., Shively J. M. and Cannon G. C. (2008). CO_2 fixation kinetics of *Halothiobacillus neapolitanus* mutant carboxysomes lacking carbonic anhydrase suggest the shell acts as a diffusional barrier for CO_2. *The Journal of Biological Chemistry,* **283 (16),** 10377–10384. DOI:10.1074/jbc.M709285200
83. Durán-Viseras A., Andrei A. S., Ghai R., Sánchez-Porro C. and Ventosa A. (2019). New *Halonotius* species provide genomics based insights into cobalamin synthesis in Haloarchaea. *Front. Microbiol.* **10**:1928. DOI:10.3389/fmicb.2019.01928
84. Egener T., Martin D. E., Sarkar A. and Reinhold-Hurek B. (2001). Role of a ferredoxin gene cotranscribed with the nifHDK operon in N_2 fixation and nitrogenase "Switch-Off" of *Azoarcus* sp. strain BH72. *Journal of bacteriology,* **183(12),** 3752-3760. DOI: 10.1128/JB.183.12.3752-3760.2001
85. Emanuelsson B. D., Baisden W. T., Bertler N. A. N., Keller E. D. and Gkinis V. (2015). High-resolution continuous-flow analysis setup for water isotopic measurement from ice cores using laser spectroscopy. *Atmospheric Measurement Techniques.* **8(7),** 2869-2883. DOI:10.5194/amt-8-2869-2015
86. Fahrenfeld N. L., Reyes H. D., Eramo A., Akob D. M., Mumford A. C. and Cozzarelli I. M. (2017). Shifts in microbial community structure and function in surface waters impacted by unconventional oil and gas wastewater revealed by metagenomics. *Science of The Total Environment,* **580,** 1205-1213. DOI:10.1016/j.scitotenv.2016.12.079

87. Fazzini M., Bisci C. and Billi P. (2015) The Climate of Ethiopia. In: Billi P. (eds) Landscapes and Landforms of Ethiopia. World Geomorphological Landscapes. Springer, Dordrecht. DOI:10.1007/978-94-017-8026-1_3

88. Feller G. (2010). Protein stability and enzyme activity at extreme biological temperatures. *Journal of physics. Condensed matter : an Institute of Physics journal,* **22(32),** 323101. DOI:10.1088/0953-8984/22/32/323101

89. Feller G. (2013). Psychrophilic enzymes: from folding to function and biotechnology. *Scientifica.* **2013,** 512840. DOI:10.1155/2013/512840

90. Feng, G., Xie, T., Wang, X. Bai J., Tang L., Zhao H., Wei W., Wang M. and Zhao Y. (2018). Metagenomic analysis of microbial community and function involved in cd-contaminated soil. *BMC Microbiology,* **18,** 11. DOI:10.1186/s12866-018-1152-5

91. Ferreira A. C, Nobre M. F., Rainey F. A., Silva M. T., Wait R., Burghardt J., Chung A. P. and Da Costa M, S. (1997). *Deinococcus geothermalis sp. nov.* and *Deinococcus murrayi sp. nov.,* two extremely radiation-resistant and slightly t*hermophilic* species from hot springs. *International Journal of Systematic Bacteriology,* **47(4),** 939–947. DOI:10.1099/00207713-47-4-939

92. Ferreira A., Nobre M., Moore E., Rainey F. A. , Battista J. R. and da Costa M.S. (1999). Characterization and radiation resistance of new isolates of *Rubrobacter radiotolerans* and *Rubrobacter xylanophilus*. *Extremophiles,* **3,** 235–238. DOI:10.1007/s007920050121

93. Filippidou S., Wunderlin T., Junier T., Jeanneret N., Dorador C., Molina V., Johnson D. R. and Junier P. (2016). A combination of extreme environmental conditions favor the prevalence of endospore forming Firmicutes. *Frontiers in Microbiology,* **7,** 1707. DOI:10.3389/fmicb.2016.01707

94. Fowler J., Toth C. R. A. and Gieg L. M. (2016). Community structure in methanogenic enrichments provides insight into syntrophic interactions in hydrocarbon-impacted environments. *Frontiers in Microbiology,* **7,** 562. DOI:10.3389/fmicb.2016.00562

95. Franzson H., Helgadóttir H. M., and Óskarsson F. (2015). Surface exploration and first conceptual model of the Dallol geothermal area, northern Afar, Ethiopia. *Proceedings World Geothermal Congress Melbourne, Australia.* 19-25 April 2015

96. Fuller M. E., Streger S. H., Rothmel R. K., Mailloux B. J., Hall J. A., Onstott T. C., Fredrickson J. K., Balkwill D. L., and DeFlaun M. F. (2000). Development of a vital fluorescent staining method for monitoring bacterial transport in subsurface environments. *Applied and environmental microbiology,* **66(10),** 4486–4496. DOI:10.1128/AEM.66.10.4486-4496.2000

97. Gaines W. L., Harrod J. R., and Lehmkuhl J. F. (1999). Monitoring biodiversity: quantification and interpretation; General Technical Report PNW-GTR-443, USDA Forest Service, Pacific North-West Research Station. https://books.google.com.et/books?id=DJOXTIMrO6IC

98. Gasteiger E., Hoogland C., Gattiker A., Duvaud S., Wilkins M. R., Appel R. D., Bairoch A. (2005). Protein Identification and Analysis Tools on the ExPASy Server. (In) John M. Walker (ed): The Proteomics Protocols Handbook, Humana Press . pp. 571-607

99. Gebresilassie S., Gebretsadik H. T. and Kabeto K. (2011). Preliminary study on geology, mineral potential and characteristics of hot springs from Dallol area, Afar rift, northeastern Ethiopia: Implications for natural resource exploration. *Momona Ethiopian Journal of Science* , **3(2)**. DOI:10.4314/mejs.v3i2.67710

100. Geourjon C. and Deleage G. (1995). SOPMA: significant improvements in protein secondary structure prediction by consensus prediction from multiple alignments. *Computer applications in the biosciences : CABIOS,* **11(6),** 681–684. DOI:10.1093/bioinformatics/11.6.681

101. Gibson M., Wang B., Ahmadi S., Burnham C. A., Tarr P. I., Warner B. B., and Dantas G. (2016). Developmental dynamics of the preterm infant gut microbiota and antibiotic resistome. *Nature Microbiology,* **1,** 16024. DOI:10.1038/nmicrobiol.2016.24

102. Ghebrezgabher M. G., Yang T. and Yang X. (2016). Long-Term Trend of Climate Change and Drought Assessment in the Horn of Africa. *Advances in Meteorology,* **2016(2),** 1-12. DOI:10.1155/2016/8057641

103. Ghurye J. S., Cepeda-Espinoza V., and Pop M. (2016). Metagenomic assembly: overview, challenges and applications. *The Yale Journal of Biology and Medicine,* **89(3),** 353–362. PMID: 27698619

104. Go G., Lee J., Choi D., Kim S. S. and Gho Y. S. (2019). Extracellular vesicle mimetic ghost nanovesicles for delivering anti-inflammatory drugs to mitigate Gram-negative bacterial outer membrane vesicle–induced systemic inflammatory response syndrome. *Advanced Healthcare Materials,* **8(4)**, e1801082. DOI:10.1002/adhm.201801082

105. Gomez F. (2014). Extreme Environment. In:Amils R. et al. (eds) Encyclopedia of Astrobiology. Springer, Berlin, Heidelberg.

106. Gómez F., Cavalazzi B., Rodríguez N., Amils R., Ori G. G., Olsson-Francis K., Escudero C., Martínez J. M. and Hagos M. (2019). Ultra-small microorganisms in

the polyextreme conditions of the Dallol volcano, northern Afar, Ethiopia. *Scientific Reports,* **9,** 7907. DOI:10.1038/s41598-019-44440-8

107. González-Torres P. and Gabaldón T. (2018). Genome variation in the model halophilic Bacterium *Salinibacter ruber*. F*rontiers in Microbiology,* **9,** 1499. DOI:10.3389/fmicb.2018.01499

108. Goto D. K. and Yan T. (2011). Genotypic diversity of *Escherichia coli* in the water and soil of tropical watersheds in Hawaii. *Applied and Environmental Microbiology,* **77(12),** 3988–3997. DOI:10.1128/AEM.02140-10

109. Grazziotin A. L., Vidal N. M. and Venancio T. M. (2015). Uncovering major genomic features of essential genes in bacteria and a methanogenic archaea. *The FEBS Journal.* **282(17):** 3395–3411. DOI:10.1111/febs.13350

110. Gregersen L. H., Bryant D. A. and Frigaard N. (2011). Mechanisms and evolution of oxidative sulfur metabolism in green sulfur bacteria. *Frontiers in Microbiology,* **2,** 116. DOI:10.3389/fmicb.2011.00116

111. Gregor I., Dröge J., Schirmer M., Quince C. and McHardy A. C. (2016). PhyloPythiaS+: a self-training method for the rapid reconstruction of low-ranking taxonomic bins from metagenomes. *PeerJ,* **4,** e1603. DOI:10.7717/peerj.1603

112. Gu W., Miller S., and Chiu, C. Y. (2019). Clinical metagenomic Next-Generation Sequencing for pathogen detection. *Annual review of pathology,* **14**, 319–338. DOI:10.1146/annurev-pathmechdis-012418-012751

113. Guan Z., Chen Y. L., Yuan .Y, Song J., Yang D. C., Xue Y., and He Y. H. (2014) Earthworm is a versatile and sustainable biocatalyst for organic synthesis. *Plos One,* **9(8),** e105284. DOI:10.1371/journal.pone.0105284

114. Guan B., Hoef V. V., Jobava R., Elroy-Stein O., Valasek L. S., Cargnello M., Gao X., Krokowski D., Merrick W. C., Kimball S. R., Komar A. A., Koromilas A. E., Wynshaw-Boris A., Topisirovic I., Ola L. and Hatzohlou M (2017). A unique ISR program determines cellular responses to chronic stress. *Molecular Cell,* **68(5),** 885–900.e6. DOI:10.1016/j.molcel.2017.11.007

115. Hallsworth J. E., Yakimov M. M., Golyshin P. N., Gillion J. L. M., D’Auria G., Alves F., La Cono V., Genovese M., McKew B. A., Hayes S. L., Harris G., Giuliano L., Timmis K. N. and McGenity T. G. (2007). Limits of life in $MgCl_2$-containing environments: chaotropicity defines the window. *Environmental Microbiology,* **9(3),** 801–813. DOI:10.1111/j.1462-2920.2006.01212.x

116. Hamilton T. L., Lange R. K., Boyd E. S. and Peters J. W. (2011). Biological nitrogen fixation in acidic high-temperature geothermal springs in Yellowstone National Park, Wyoming. *Environmental Microbiology,* **13(8),** 2204-15. DOI:10.1111/j.1462-2920.2011.02475.x

117. Han W., He P., Lin Y., Shao L. and Lü F. (2019). A Methanogenic consortium was active and exhibited long-term survival in an extremely acidified thermophilic bioreactor. *Frontiers in Microbiology,* **10**, 2757.DOI:10.3389/fmicb.2019.02757

118. Handelsman J., Rondon M. R., Brady S. F., Clardy J. and Goodman R. M. (1998). Molecular biological access to the chemistry of unknown soil microbes: a new frontier for natural products. *Chemistry and Biology,* **5(10),** R245-R249. DOI:10.1016/s1074-5521(98)90108-9

119. Harding T., Roger A.J. and Simpson A.G.B. (2017). Adaptations to high salt in halophilic protist: differential expression and gene acquisitions through duplications

and gene transfers. *Frontiers in Microbiology,* **8,** 944. DOI:10.3389/fmicb.2017.00944

120. Hbarbalz F. M., Orellana E., Figuerola E. L. M. and Erijman L. (2016). Shotgun metagenomic profiles have a high capacity to discriminate samples of activated sludge according to wastewater type. *Applied and Environmental Microbiology,* **82(17),** 5186-5196. DOI:10.1128/AEM.00916-16
121. Hedlund B. P., Mcdonald A. I., Lam J., Dodsworth J. A., Brown J. R. and Hungate B. A. (2011). Potential role of *Thermus thermophilus* and *T. oshimai* in high rates of nitrous oxide (N_2O) production in 80°C hot springs in the US Great Basin. *Geobiology,* **9,** 471-480. DOI:10.1111/j.1472-4669.2011.00295.x
122. Herbstein F.H., Kapon M. and Weissman A. (1982). X-ray diffraction as a tool for studying stoichiometry and kinetics of solid state thermal decomposition reactions: Application to the thermal decomposition of bischofite $MgCl_2 \cdot 6H_2O$. Israel *Journal of Chemistry,* **22(3),** 207-213. DOI:10.1002/IJCH.198200041
123. Herrmann I. K., Wood M. J. A. and Fuhrmann G. (2021). Extracellular vesicles as a next-generation drug delivery platform. *Nature Nanotechnology,* **16,** 748–759. DOI:10.1038/s41565-021-00931-2
124. Hernandez J. A., Curatti L., Aznar C. P., Perova Z., Britt R. D., and Rubio L. M. (2008). Metal trafficking for nitrogen fixation: NifQ donates molybdenum to NifEN/NifH for the biosynthesis of the nitrogenase FeMo-cofactor. *Proceedings of the National Academy of Sciences of the United States of America,* **105(33),** 11679–11684. DOI: 10.1073/pnas.0803576105

125. Hesse S. J. A., Ruijter G. J. G., Dijkema C. and Visser J. (2002), Intracellular pH homeostasis in the filamentous fungus *Aspergillus niger. European Journal of Biochemistry,* **269(14),** 3485-3494. DOI:10.1046/j.1432-1033.2002.03042.x

126. Hill G. T., Mitkowskia N. A., Aldrich-Wolfeb L., Emelea L. R., Jurkoniea D. D., Fickea A., Maldonado-Ramireza S., Lyncha S. T. and Nelsona E. B. (2000). Methods for assessing the composition and diversity of soil microbial communities. *Applied Soil Ecology,* **15(1),** 25–36. DOI:10.1016/S0929-1393(00)00069-X

127. Hollien J. and Marqusee S. (1999). Structural distribution of stability in thermophilic enzyme. *Proceedings of the National Academy of Sciences of the United States of America,* **96(24),** 13674–13678. DOI:10.1073/pnas.96.24.13674

128. Hoque E., Pflugmacher,S., Fritscher J. and Wolf M. (2007). Induction of glutathione S-transferase in biofilms and germinating spores of *Mucor hiemalis* strain EH5 from cold sulfidic spring waters. A*pplied and environmental microbiology,* **73(8),** 2697–2707. DOI:10.1128/AEM.02786-06

129. Hosseinkhani B., van den Akker N., D'Haen J., Gagliardi M., Struys T., Lambrichts I., Waltenberger J., Nelissen I., Hooyberghs J., Molin D. G. M. and Michiels L. (2017). Direct detection of nano-scale extracellular vesicles derived from inflammation-triggered endothelial cells using surface plasmon resonance. *Nanomedicine,* **13(5),** 1663-1671. DOI: 10.1016/j.nano.2017.03.010

130. Hou W., Wang S., Dong H., Jiang H., Briggs B. R., Peacock J. P., Huang Q., Huang L., Wu G., Zhi X., Li W., Dodsworth J. A., Hedlund B. P., Zhang C., Hartnett H. E., Dijkstra P. and Hungate B. A. (2013). A comprehensive census of microbial diversity in hot springs of Tengchong, Yunnan Province China using 16S rRNA gene pyrosequencing. *Plos One,* **8(1),** e53350. DOI:10.1371/journal.pone.0053350

131. Howe A. C., Jansson J. K., Malfatti S. A., Tringe S. G., Tiedje J. M., Brown C. T. (2014). Tackling soil diversity with the assembly. *Proceedings of the National Academy of Sciences of the United States of America,* **111 (13),** 4904-4909. DOI: 10.1073/pnas.1402564111

132. Huang W., Ertekin I., Wang T., Cruz L., Dailey M., DiRuggiero J. and Kisailus D. (2020). Mechanism of water extraction from gypsum rock by desert colonizing microorganisms. *Proceedings of the National Academy of Sciences,* **117(20),** 10681-10687. DOI: 10.1073/pnas.2001613117

133. Hugenholtz P., Goebel B. M. and Pace N. R. (1998). Impact of culture-independent studies on the emerging phylogenetic view of bacterial diversity. *Journal of Bacteriology,* **180(18),** 4765-74. DOI: 10.1128/JB.180.18.4765-4774.1998

134. Hughes B. S., Cullum A. J. and Bennett A. F. (2007), Evolutionary adaptation to environmental pH in experimental lineage of *Escherichia coli. Evolution,* **61,** 1725-1734. DOI:10.1111/j.1558-5646.2007.00139.x

135. Hughes R.cH. and Hughes J.cS. (1992). A directory of African wetlands. The World Conservation Union (IUCN), Cambridge UK, The United Nations Environmental Programme (UNEP), Nairobi Kenya, The World Conservation Monitoring Center (WCMC), Cambrifge, UK, 167-168.

136. Huson D., Beier S., Flade I., Gorska A., El-Hadidi M., Mitra S., Ruscheweyh H. J. and Tappu R. (2016). MEGAN community edition - interactive exploration and analysis of large-scale microbiome sequencing data. *Plos Computational Biology,* **12(6),** e1004957. DOI:10.1371/journal. Pcbi.1004957

137. Huson D. H., Albrecht B., Bağcı C., Bessarab I., Górska A., Jolic D. and Williams R. B. H. (2018). MEGAN-LR: new algorithms allow accurate binning and easy

interactive exploration of metagenomic long reads and contigs. *Biology Direct,* **13(1),** 6. DOI: 10.1186/s13062-018-0208-7

138. Ibarbalz F. M., Orellana E., Figuerola E. L. M and Erijman L. (2016). Shotgun metagenomic profiles have a high capacity to discriminate samples of activated sludge according to wastewater type. *Applied and Environmental Microbiology*. **82(17),** 5186–5196. DOI:10.1128/AEM.00916-16
139. Illsley-Kemp F., Keir D., Bull J. M., Ayele A., Hammond J. O. S., Kendall J. M., Gallacher R. J., Gernon T. and Goitom B. (2017). Local earthquake magnitude scale and b-Value for the Danakil region of northern Afar. *Bulletin of the Seismological Society of America,* **107(2),** 521-531. DOI:10.1785/0120150253
140. Imperial J., Ugalde R. A., Shah V. K. and Brill W. J. (1984). Role of the nifQ gene product in the incorporation of molybdenum into nitrogenase in *Klebsiella pneumoniae. Journal of Bacteriology*. **158(1),** 187-94. DOI:10.1128/jb.158.1.187-194.1984
141. Isaji Y., Kawahata H., Ogawa N. O., Kuroda J., Yoshimura T., Jiménez-Espejo F. J., Makabe A., Shibuya T., Lugli S., Santulli A., Manzi V., Roveri M. and N Ohkouchi. (2019). Efficient recycling of nutrients in modern and past hypersaline environments. *Scientific Reports,* **9,** 3718 . DOI:10.1038/s41598-019-40174-9
142. Ishii S. and Sadowsky M.J. (2008) *Escherichia coli* in the environment: implications for water quality and human health. *Microbes and environments*, **23(2),** 101–108. DOI:10.1264/jsme2.23.101
143. Ivleva N. B., Groat J., Staub J. M. and Stephens M. (2016). Expression of active subunit of nitrogenase via integration into plant organelle genome. *Plos One,* **11(8),** e0160951. DOI: 10.1371/journal.pone.0160951

144. Jacob H. J., Hussein E. I., Shakhatreh M. A. K. and Corn C. T. (2017). Microbial community analysis of the hypersaline water of the Dead Sea using high-throughput amplicon sequencing. *MicrobiologyOpen*, **6(5),** e00500. DOI: 10.1002/mbo3.500
145. Jahromi L. P. and Fuhrmann G. (2021). Bacterial extracellular vesicles: understanding biology promotes applications as nanopharmaceuticals. *Advanced Drug Delivery Reviews*, **173,** 125-140. DOI: 10.1016/j.addr.2021.03.012
146. Jay E., Bambara R., Padmanabhan R. and Wu R. (1974). DNA sequence analysis: a general, simple and rapid method for sequencing large oligodeoxyribonucleotide fragments by mapping. *Nucleic Acids Research,* **1(3),** 331–354. DOI:10.1093/nar/1.3.331
147. Jiang H., Dong H., Zhang G., Yu B., Chapman L. R. and Fields M. W. (2006). Microbial diversity in water and sediment of Lake Chaka, an athalassohaline lake in northwestern China. *Applied and Environmental Microbiology,* **72(6),** 3832–3845. DOI:10.1128/AEM.02869-05
148. Jimenez D. J., Andreote F. D., Chaves D., Montana J. S., Osorio-Forero C., Junca H., Zambrano M. M. and Baena S. (2012). Structural and functional insights from the metagenome of an acidic hot spring microbial plancktonic community in the Colombian Andes. *Plos One,* **7(12),** e52069. DOI:10.1371/journal.pone.0052069
149. Jimenez D. J. and Zambrano M. M. (2015). Metagenome of acidic hot spring microbial planktonic community: structural and functional insights. In: Nelson K.E. (eds) Encyclopedia of Metagenomics. Springer, Boston, MA. DOI:10.1007/978-1-4899-7478-5_173
150. Jimenez-Vicente E., Yang Z., Martin Del Campo J. S., Cash V. L., Seefeldt L. C. and Dean D. R. (2019). The NifZ accessory protein has an equivalent function in

maturation of both nitrogenase MoFe protein P-clusters. *The Journal of Biological Chemistry,* **294(16),** 6204–6213. DOI: 10.1074/jbc.RA119.007905

151. Jimenez-Vicente E., Yang Z., Ray W. K., Echavarri-Erasun C., Cash V. L., Rubio L. M., Seefeldt L. C. and Dean D. R. (2018). Sequential and differential interaction of assembly factors during nitrogenase MoFe protein maturation. *The Journal of Biological Chemistry,* **293(25),** 9812–9823. DOI: 10.1074/jbc.RA118.002994
152. Jin Y., Zhou J., Zhou J., Hu M., Zhang Q., Kong N., Ren H., Liang L. and Yue J. (2020). Genome-based classification of *Burkholderia cepacia* complex provides new insight into its taxonomic status. *Biology Direct,* **15(1),** 6. DOI:10.1186/s13062-020-0258-5
153. Johnson D. B., Rolfe S., Hallberg K. B. and Iversen E. (2001). Isolation and phylogenetic characterization of acidophilic microorganisms indigenous to acidic drainage waters at an abandoned Norwegian copper mine. *Environmental Microbiology,* **3(10),** 630–637. DOI: 10.1046/j.1462-2920.2001.00234.x
154. Jones P. F., Clark M. I., King R., Shaw J. L., Woodward J. M. and Hirsch R. P. (2016). Novel European free-living, non-diazotrophic *Bradyrhizobium* isolates from contrasting soils that lack nodulation and nitrogen fixation genes- a genome comparison. *Scientific Reports,* **6,** 25858. DOI:10.1038/srep25858
155. Kanehisa M. and Sato Y. (2019). KEGG Mapper for inferring cellular functions from protein sequences. *Protein Science,* **29(1),** 28–35. DOI:10.1002/pro.3711
156. Kanehisa M. and Goto S. (2000). KEGG: kyoto encyclopedia of genes and genomes. *Nucleic acids Research,* **28(1),** 27–30. DOI:10.1093/nar/28.1.27
157. Kaur C., Selvakumar G. and Ganeshamurthy A. N. (2017). *Burkholderia* to *Paraburkholderia*: The Journey of a plant-beneficial-environmental bacterium. In:

Shukla P. (eds) Recent advances in Applied Microbiology. Springer, Singapore. DOI:10.1007/978-981-10-5275-0_10

158. Kawano Y., Suzuki K. and Ohtsu I. (2018). Current understanding of sulfur assimilation metabolism to biosynthesize L-cysteine and recent progress of its fermentative overproduction in microorganisms. *Applied Microbiology and Biotechnology,* **102(19),** 8203–8211. DOI:10.1007/s00253-018-9246-4
159. Kebede E., Mariam Z. G., and Ahlgren I. (1994). The Ethiopian Rift Valley lakes: chemical characteristics of a salinity-alkalinity series. *Hydrobiology,* **288(1),** 1-12. DOI:10.1007/bf00006801
160. Ketchum R. N., Smith E. G., Vaughan G. O., Phippen B. L., McParland D., Al-Mansoori N., Carrier T. J., Burt J. A. and Reitzel A. M. (2018). DNA extraction method plays a significant role when defining bacterial community composition in the marine invertebrate *Echinometra mathaei. Frontiers in Marine Science,* **5,** 255. DOI:10.3389/fmars.2018.00255
161. Kim B. H. and Gadd G. M. (2008). Degradation of nucleic acid bases. Bacterial physiology and metabolism (pp135- 223). New York. Cambridge University Press.
162. Kim H. S., Lee O. K., Lee S. J., Hwang S., Kim S. J., Yang S., Park S. and Lee E. Y. (2006). Enantioselective epoxide hydrolase activity of a newly isolated microorganism, *Sphingomonas echinoides EH-983,* from seawater. *Journal of Molecular Catalysis B: Enzymatic,* **41(3–4),** 130-135. DOI:10.1016/j.molcatb.2006.05.009
163. Kimura H., Nashimoto H., Shimizu M., Hattori S., Yamada K., Koba K., Yoshida N. and Kato K. (2010). Microbial methane production in deep aquifer associated with

the accretionary prism in southwest Japan. *The ISME Journal*, **4(4),** 531–541. DOI:10.1038/ismej.2009.132

164. Kiyasu C., Asakura A., Nagahashi Y. and Hoshino T. (2000). Contribution of cysteine desulfurase (NifS Protein) to the biotin synthase reaction of *Escherichia coli. Journal of Bacteriology,* **182(10),** 2879-2885. DOI: 10.1128/JB.182.10.2879-2885.2000
165. Kleiner M., Thorson E., Sharp C. E. Dong X., Liu D., Li C. and Strous M. (2017). Assessing species biomass contributions in microbial communities via metaproteomics. *Nature communications,* **8(1),** 1558. DOI:10.1038/s41467-017-01544-x
166. Kluber L. A., Johnston E. R., Allen S. A., Hendershot J. N., Hanson P. J. and Schadt C. W. (2020). Constraints on microbial communities, decomposition and methane production in deep peat deposits. *Plos One,* **15(2),** e0223744. DOI:10.1371/journal.pone.0223744
167. Kodzius R. and Gojobori T. (2015). Marine metagenomics as a source for bioprospecting. *Marine Genomics,* **24(1),** 21-30. DOI:10.1016/j.margen.2015.07.001
168. Koonin E. V. and Galperin M. Y. (2003). Sequence evolution function: computational approaches in comparative genomics. Boston: Kluwer Academic. Chapter 4, Principles and Methods of Sequence Analysis. PMID: 21089240
169. Koslicki D., Foucart S. and Rosen G. (2014). WGSQuikr: Fast whole-genome shotgun metagenomic classification. *Plos One,* **9(3),** e91784. DOI:10.1371/journal.pone.0091784
170. Kotopoulou E., Huertas A.D., Garcia-Ruiz J.M., Dominguez-Vera J.M., Lopez-Garcia J.M., Guerra-Tschuschke I. and Rull F. (2019).A polyextreme hydrothermal

system controlled by iron: The case of Dallol at the Afar Triangle. *ACS Earth and Space Chemistry,* **3(1),** 90–99. DOI:10.1021/acsearthspacechem.8b00141

171. Kouza M., Faraggi, E., Kolinski, A., and Kloczkowski, A. (2017). The GOR method of protein secondary structure prediction and its application as a protein aggregation prediction tool. *Methods in Molecular Biology*, **1484,** 7–24. DOI:10.1007/978-1-4939-6406-2_2
172. Kurade M. B., Saha S., Salama E., Patil S. M., Govindwar S. P. and Jeon B. (2019). Acetoclastic methanogenesis led by Meth-anosarcina in anaerobic co-digestion of fats, oil and grease for enhanced production of methane. *Bioresource Technology*, **272,** 351–359. DOI: 10.1016/j.biortech.2018.10.047
173. Kurth J. M., Op den Camp H. J. M. and Welte C. U. (2020). Several ways one goal methanogenesis from unconventional substrates. *Applied Microbiology and Biotechnology,* **104(16),** 6839–6854. DOI:10.1007/s00253-020-10724-7
174. Kushkevych I., Cejnar J., Treml J., Dordević D., Kollar P. and Vítězová M. (2020). Recent advances in metabolic pathways of sulfate reduction in intestinal bacteria. *Cells,* **9(3),** 698. DOI:10.3390/cells9030698
175. Lalou C., Van N. H., Faure H., and Moreira L. (1970). Datation by uranium-thorium method of high coral reefs from Afar tectonic depression (Ethiopia). *Revue de Geographie Physique et de Geologie Dynamique,* **12(1),** 3-8.
176. Lan Y., Rosen G. and Hershberg R. (2016). Marker genes that are less conserved in their sequences are useful for predicting genome-wide similarity levels between closely related prokaryotic strains. *Microbiome,* **4(1),** 18. DOI:10.1186/s40168-016-0162-5

177. Lanzén A., Simachew A., Gessesse A., Chmolowska D., Jonassen I., and Øvreås L. (2013). Surprising prokaryotic and eukaryotic diversity, community structure and biogeography of Ethiopian soda lakes. *Plos One,* **8(8),** e72577. DOI:10.1371/journal.pone.0072577

178. Lemaire O. N., Jespersen M., and Wagner T. (2020). CO_2-fixation strategies in energy extremophiles: What can we learn from Acetogens?. *Frontiers in microbiology,* **11,** 486. DOI: 10.3389/fmicb.2020.00486

179. Lewis K. (2009). Persisters, biofilms, and the problem of cultivability. In S Epstein (Ed). Uncultivated Microorganisms. (pp. 181–94). Berlin: Springer-Verlag.

180. Lesser M. P., Morrow K. M., Pankey S. M. and Noonan S. H. C. (2018). Diazotroph diversity and nitrogen fixation in the coral *Stylophora pistillata* from the Great Barrier Reef. *The ISME Journal,* **12(3),** 813–824. DOI:10.1038/s41396-017-0008-6

181. Li S. J., Hua Z. S., Huang L. N., Shi S. H., Chen L. X., Kuang J. L., Liu J., Hu M. and Shu W. S. (2014). Microbial communities evolve faster in extreme environments. *Scientific Reports,* **4,** 6205. DOI: 10. 1038/srep06205

182. Li X., Liu Q., Liu X., Shi H. and Chen S. (2016). Using synthetic biology to increase nitrogenase activity. *Microbial cell factories,* **15,** 43. DOI:10.1186/s12934-016-0442-6

183. Li Y., Tang K., Zhang L., Zhao Z., Xie X., Chen C. A., Wang D., Jiao N. and Zhang Y. (2018). Coupled carbon, sulfur, and nitrogen cycles mediated by microorganisms in the water column of a shallow-water hydrothermal ecosystem. *Frontiers in Microbiology,* **9,** 2718. DOI:10.3389/fmicb.2018.02718

184. Lindgreen S., Adair K. and Gardner P. (2016). An evaluation of the accuracy and speed of metagenome analysis tools. *Scientific Reports,* **6,** 19233. DOI:10.1038/srep19233

185. Littlechild J. (2015). Enzymes from extreme environments and their industrial applications. *Frontiers in Bioengineering and Biotechnology,* **3,** 161. DOI:10.3389/fbioe.2015.00161

186. Liu B., Faller L. L., Klitgord N., Mazumdar V., Ghodsi M., Sommer D. D., Gibbons T. R., Treangen T. J., Chang Y. C., Li S., Stine O. C., Hasturk H., Kasif S., Segrè D., Pop M., and Amar S. (2012^{a}). Deep sequencing of the oral microbiome reveals signatures of periodontal disease. *Plos One,* **7(6),** e37919. DOI:10.1371/journal.pone.0037919

187. Liu L., Li Y., Li S., Hu N., He Y., Lin P.D., Lu L. and Law M. (2012^{b}). Comparison of Next-Generation sequencing Systems. *Journal of Biomedicine and Biotechnology,* **2012,** 251364. DOI:10.1155/2012/251364

188. Liu Y. X., Qin Y., Chen T., Lu M., Qian X., Guo X. and Bai Y. (2020^{a}). A practical guide to amplicon and metagenomic analysis of microbiome data. *Protein and Cell,* **12(5),** 315–330. DOI:10.1007/s13238-020-00724-8

189. Liu Z., Wang K., Chen Y., Tan T. and Nielsen J. (2020^{b}). Third-generation biorefineries as the means to produce fuels and chemicals from CO2. *Nature Catalysis,* **3,** 274–288. DOI:10.1038/s41929-019-0421-5

190. Logares R., Lindström E., Langenheder S., Logue J. B., Paterson H, Laybourn-Parry J., Rengefors K., Tranvik L. and Bertilsson S. (2013). Biogeography of bacterial communities exposed to progressive long-term environmental change. *The ISME Journal,* **7,** 937–948. DOI:10.1038/ismej.2012.168

191. López-García J. M., Moreira D., Benzerara K., Grunewald O., López-García P. (2020). Origin and evolution of the halo-volcanic complex of Dallol: Proto-volcanism in northern Afar (Ethiopia). *Frontiers in Earth Science,* **7,** 351. DOI:10.3389/feart.2019.00351
192. Luk K., C. Berg M. G., Naccache S. N., Kabre B., Federman S., Mbanya D., Kaptué L., Chiu C. Y., Brennan C. A. and Hackett J. J. (2015). Utility of metagenomic next-generation sequencing for characterization of HIV and human pegivirus diversity. *Plos One,* **10,** e0141723. DOI:10.1371/journal.pone.0141723
193. Luo C., Tsementzi D., Kyrpides N., Read T. and Konstantinidis K. T. (2012). Direct comparisons of Illumina vs. Roche 454 sequencing technologies on the same microbial community DNA sample. *Plos One,* **7(3),** e30087. DOI: 10.1371/journal.pone.0030087
194. Makhdoumi-Kakhki A., Amoozegar M. A., Kazemi B., PaiC L., and Ventosa A. (2012). Prokaryotic diversity in Aran-Bidgol salt lake, the largest hypersaline playa in Iran. *Microbes and Environments,* **27(1),** 87-93. DOI: 10.1264/jsme2.me11267
195. Mamo G. (2020). Challenges and adaptations of life in alkaline habitats. *Advances in Biochemical Engineering/Biotechnology,* **172,** 85-133. DOI:10.1007/10_2019_97
196. Martin T., Ifo S. A., Moutsambote J., Koubouana F., Yoka J., Ndzai S. F., Bouetou-Kadilamio L. N. O., Mampouya H., Jourdain C., Bocko Y., Mantota A. B., Mbemba M., Mouanga-Sokath D., Odende R., Mondzali L. R., Wenina Y. E. M. Ouissika B. C. and Joel L. J. (2016). Tree species diversity, richness, and similarity in intact and degraded forest in the tropical rain forest of the Congo Basin: Case of the forest of Likouala in the Republic of Congo. *International Journal of Forestry Research,* **1687,** 9368. DOI:10.1155/2016/7593681

197. Master S. (2016). Gaet'ale- a reactivated thermal spring and potential tourist hazard in the Asale salt flats, Danakil Depression, Ethiopia. *Journal of Applied Volcanology,* **5,** 1. DOI:10.1186/s13617-015-0042-x

198. Matsumoto T., Takahashi S., Kaieda M., Ueda M., Tanaka A., Fukuda H. and Kondo A. (2001). Yeast whole-cell biocatalyst constructed by intracellular overproduction of Rhizopus oryzae lipase is applicable to biodiesel fuel production. *Applied Microbiology and Biotechnology,* **57,** 515–520. DOI:10.1007/s002530100733

199. Mayer K., Scheu B., Yilmaz T. I., Montanaro C., Gilg H. A., Rott S., Joseph E. P. and Dingwell D. B. (2017). Phreatic activity and hydrothermal alteration in the Valley of Desolation, Dominica, Lesser Antilles. *Bulletin of Volcanology,* **79(12),** 82. DOI:10.1007/s00445-017-1166-0

200. McCully A. Onyeziri M., LaSarre B., Gliessman J. and McKinlay J. (2019). Reductive tricarboxylic acid cycle enzymes and reductive amino acid synthesis pathways contribute to electron balance in a *Rhodospirillum rubrum* Calvin-cycle mutant. *Microbiology,* **166,** DOI:10.1099/mic.0.000877

201. McGenity T. J. and Sorokin D. Y. (2019). Methanogens and methanogenesis in hypersaline environment. In Stams A., Sousa D. (eds) Biogenesis of hydrocarbons. Handbook of Hydrocarbon and Lipid Microbiology. pp 283-309. Springer, Cham. DOI: 10.1007/978-3-319-78108-2_12

202. McGlynn S. E., Boyd E. S., Peters J. W., and Orphan V. J. (2013). Classifying the metal dependence of uncharacterized nitrogenases. *Frontiers in microbiology,* **3,** 419. DOI:10.3389/fmicb.2012.00419

203. McHardy A. C., Hector G. M., Aristotelis T., Philip H. and Isidore R. (2007). Accurate phylogenetic classification of variable-length DNA fragments. *Nature Methods,* **4(1),** 63–72. DOI:10.1038/nmeth976

204. Medeiros J. D., Cantão M. E., Cesar D. E., Nicolás M. F., Diniz C. G., Silva V. L., de Vasconcelos A. T. R. and Coelho C. M. (2016). Comparative metagenome of a stream impacted by the urbanization phenomenon. *Brazilian Journal of Microbiology,* **47(4),** 835-845. DOI:10.1016/j.bjm.2016.06.011

205. Mee M. T., Collins J. J., Church G. M. and Wang H. H. (2014). Syntrophic exchange in synthetic microbial communities. *Proceedings of the National Academy of Sciences of the United States of America,* **111(20),** E2149–E2156. DOI:10.1073/pnas.1405641111

206. Meena K. K., Sorty A. M., Bitla U. M., Choudhary K., Gupta P., Pareek A., Singh D. P., Prabha R., Sahu P. K., Gupta V. K., Singh H. B., Krishanani K. K. and Minhas P. S. (2017). Abiotic stress responses and microbe-mediated mitigation in plants: The Omics strategies. *Frontiers in Plant Science,* **8,** 172. DOI:10.3389/fpls.2017.00172

207. Meiser A., Otte J., Schmitt I. and Grande F. D. (2017). Sequencing genomes from mixed DNA samples- evaluating the metagenome skimming approach in lichenized fungi. *Scientific Report,* **7,** 14881. DOI:10.1038/s41598-017-14576-6

208. Mendes L. W., Braga L. P. P., Navarrete A. A., de Souza D. G., Silva G. G. Z. and Tsai S. M. (2017). Using metagenomics to connect microbial community biodiversity and functions. *Current Issues in Molecular Biology,* **24,** 103-118. DOI:10.21775/cimb.024.103

209. Mendez-Garcia C., Pelaez A. I., Mesa V., Sanchez J., Golyshina O. V. and Ferrer M. (2015). Microbial diversity and metabolic networks in acid mine drainage habitats. *Frontiers in Microbiology,* **6,** 475. DOI: 10.3389/fmicb.2015.00475

210. Meng L., Li W., Zhang S., Wu C. and Wang K. (2016). Effects of sucrose amendment on ammonia assimilation during sewage sludge composting. *Bioresource Technology,* **210,** 160-166. DOI:10.1016/j.biortech.2016.01.094

211. Merino N., Aronson H. S., Bojanova D. P., Feyhl-Buska J., Wong M. L., Zhang S., and Giovannelli D. (2019). Living at the extremes: Extremophiles and the limits of life in a planetary context. *Frontiers in microbiology,* **10,** 780. DOI:10.3389/fmicb.2019.00780

212. Mernagh T. P., Bastrakov E. N., Clarke J. D. A., de Caritat P., Dulfer H., English P. M., Jaireth S., Thomas M. and Trafford J. (2014). Lake Tyrrell, Australia, and its potential for strategic resources. Promotional brochure. Geoscience, Australia, Canberra. 6p.

213. Mesa V., Gallego J. L. R., Gonzalez-Gil R., Lauga B., Sanchez J., Mendez-Garcia C. and Pelaez A. I. (2017). Bacterial, archaeal, and eukaryotic diversity across distinct microhabitats in an acid mine drainage. *Frontiers in Microbiology,* **8,** 1756. DOI: 10.3389/fmicb.2017.01756

214. Messer L., Mahaffey C., Robinson C. M., Jeffries T. C., Baker K. G., Isaksson J. B., Ostrowski M., Doblin M. A., Brown M. V. and Seymour J. R. (2016). High levels of heterogeneity in diazotroph diversity and activity within a putative hotspot for marine nitrogen fixation. *The ISME Journal,* **10,** 1499–1513. DOI:10.1038/ismej.2015.205

215. Meyer F., Paarmann D., D'Souza M. Olson R., Glass E. M., Kubal M., Paczian T., Rodriguez A., Stevens R., A Wilke, Wilkening J. and Edwards R. A. (2008). The metagenomics RAST server – a public resource for the automatic phylogenetic and functional analysis of metagenomes. *BMC Bioinformatics,* **9,** 386. DOI:10.1186/1471-2105-9-386

216. Meziti A., Nikouli E., Hatt J. K., Konstantinidis K. T. and Kormas K. A. (2021). Time series metagenomic sampling of the Thermopyles, Greece, geothermal springs reveals stable microbial communities dominated by novel sulfur-oxidizing chemoautotrophs. *Environmental Microbiology,* **23,** 3710-3726. DOI:10.1111/1462-2920.15373

217. Mikheenko A., Saveliev V., and Gurevich A. (2016). MetaQUAST: evaluation of metagenome assemblies. *Bioinformatics,* **32(7),** 1088–1090. DOI:10.1093/bioinformatics/btv697

218. Miramontes P. (1992). A cellular automaton model for the evolution of nucleic acids. PhD Thesis. University of Namibia.

219. Mitchell K. R., and Takacs-Vesbach C. D. (2008). A comparison of methods for total community DNA preservation and extraction from various thermal environments. *Journal of Industrial Microbiology and Biotechnology,* **35(10),** 1139-1147. DOI: 10.1007/s10295-008-0393-y

220. Momen-Heravi F., Balaj L., Alian S., Mantel P. Y., Halleck A. E., Trachtenberg A. J., Soria C. E., Oquin S., Bonebreak C. M., Saracoglu E., Skog J., and Kuo W. P. (2013). Current methods for the isolation of extracellular vesicles. Biological chemistry, **394(10),** 1253-62. DOI:10.1515/hsz-2013-0141

221. Morales P., Mencher A., Tronchoni J., Gonzalez R. (2021). Extracellular vesicles in food biotechnology. *Microbial Biotechnology,* **14(1),** 1751-7915. DOI: 10.1111/1751-7915.13657

222. Morett E., Fischer H. M. and Hennecke H. (1991). Influence of oxygen on DNA binding, positive control, and stability of the *Bradyrhizobium japonicum* NifA regulatory protein. *Journal of bacteriology,* **173(11),** 3478-3487. DOI: 10.1128/jb.173.11.3478-3487.1991

223. Mormile M. R., Hong B. Y., and Benison K. C. (2009). Molecular analysis of the microbial communities of Mars analog lakes in Western Australia. *Astrobiology,* **9(10),** 919-930. DOI: 10.1089/ast.2008.0293

224. Mozaheb N. and Mingeot-Leclercq M. P. (2020). Membrane vesicle production as a bacterial defense against stress. *Frontiers in Microbiology,* **11,** 600221. DOI: 10.3389/fmicb.2020.600221

225. Narasingarao P., Podell S., Ugalde J. A., Brochier-Armanet C., Emerson J. B., Brocks J. J., Heidelberg K. B., Banfield J. F. and Allen E. E. (2012). De novo metagenomic assembly reveals abundant novel major lineage of Archaea in hypersaline microbial communities. *The ISME Journal,* **6,** 81-93. DOI: 10.1038/ismej.2011.78

226. Nayak D. D., Liu A., Agrawal N., Rodriguez-Carerro R., Dong S. H., Mitchell D. A., Nair S. K. and Metcalf W. M. (2020) Functional interactions between posttranslationally modified amino acids of methyl-coenzyme M reductase in *Methanosarcina acetivorans. Plos Biology,* **18(2),** e3000507. DOI:10.1371/journal.pbio.3000507

227. Neelakanta G. and Sultana H. (2013). The use of metagenomic approaches to analyze changes in microbial communities. *Microbiology Insights.* **6,** 37–48. DOI:10.4137/MBI.S10819

228. Nissenbaum A. (1975). The microbiology and biogeochemistry of the Dead Sea. *Microbial Ecology,* **2(2),** 139-161. DOI: 10.1007/BF02010435

229. Nobile A. Pagli C., Keir D., Wright T.J ., Ayele A., Ruch J., Acocella A. (2012). Dike-fault interaction during the 2004 Dallol intrusion at the northern edge of the Erta Ale Ridge (Afar, Ethiopia). *Geophysical Research Letters*, **39,** L19305 DOI:10.1029/2012GL053152

230. Nonaka A., Yamamoto H., Kamiya N., Kotani H., Yamakawa H., Tsujimoto R., and Fujita Y. (2019). Accessory proteins of the nitrogenase assembly, NifW, NifX/NafY, and NifZ, are essential for Diazotrophic growth in the non heterocystous *Cyanobacterium Leptolyngbya boryana*. *Frontiers in microbiology,* **10,** 495. DOI:10.3389/fmicb.2019.00495

231. Nurk S., Meleshko D., Korobeynikov A., and Pevzner P. A. (2017). metaSPAdes: a new versatile metagenomic assembler. *Genome Research,* **27(5),** 824–834. DOI:10.1101/gr.213959.116

232. Nzila A., Thukair A., Sankara S., Chanbasha B. and Musa M. M. (2016). Isolation and characterization of naphthalene biodegrading *Methylobacterium radiotolerans* bacterium from the eastern coastline of the Kingdom of Saudi Arabia. *Archives of Environmental Protection,* **42(3),** 25-32. DOI:10.1515/aep-2016-0028

233. Orellana R., Macaya C., Bravo G., Dorochesi F., Cumsille A., Valencia R., Rojas C. and Seeger M. (2018). Living at the frontiers of life: Extremophiles in Chile and

their potential for bioremediation. *Frontiers in Microbiology,* **9,** 2309. DOI:10.3389/fmicb.2018.02309

234. Oren A. (2005). A hundred years of *Dunaliella* research: 1905–2005. *Saline Systems,* **1,** 2. DOI:10.1186/1746-1448-1-2
235. Oren A. (2006). Life at high salt concentrations. *Prokaryotes,* **2,** 263–282. DOI: 10.1007/0-387-30742-7_9
236. Oren A. (2010). The dying Dead Sea: The microbiology of an increasingly extreme environment. *Lakes and Reservoirs: Research and Management,* **15(3),** 1-8. DOI: 10.1111/j.1440-1770.2010.00435.x
237. Oren A., Gurevich P., Anati D. A., Barkan E. and Luz B. (1995). A bloom of *Dunaliella parva* in the Dead Sea in 1992: Biological and biogeochemical aspects. *Hydrobiologia,* **297 (1),** 73 -185. DOI:10.1007/BF00019283
238. Osborne P., Hall L. J., Kronfeld-Schor N., Thybert D. and Haerty W. (2020). A rather dry subject; investigating the study of arid-associated microbial communities. *Environmental Microbiome,* **15,** 20. DOI:10.1186/s40793-020-00367-6
239. Oulas A., Pavloudi C., Polymenakou P., Pavlopoulos G. A., Papanikolaou N., Kotoulas G., Arvanitidis C. and Iliopoulos I. (2015). Metagenomics: tools and insights for analyzing Next-Generation Sequencing data derived from biodiversity studies. *Bioinformatics and Biology Insights,* **9,** 75–88. DOI:10.4137/BBI.S12462
240. Overbeek R., Begley T., Butler R. M., Choudhuri J. V., Chuang H. Y., Cohoon M., de Crécy-Lagard V., Diaz N., Disz T., Edwards R., Fonstein M., Frank E. D., Gerdes S., Glass E. M., Goesmann A., Hanson A., Iwata-Reuyl D., Jensen R., Jamshidi N., Krause L., Kubal M., Larsen N., Linke B., McHardy A. C., Meyer F., Neuweger H., Olsen G., Olson R., Osterman A., Portnoy V., Pusch G. D., Rodionov D. A., Rückert

C., Steiner J., Stevens R., Thiele I., Vassieva O., Ye Y., Zagnitko O. and Vonstein V. (2005). The subsystems approach to genome annotation and its use in the project to annotate 1000 genomes. *Nucleic Acids Research,* **33(17),** 5691-702. DOI: 10.1093/nar/gki866

241. Overbeek R., Olson R., Pusch G. D., Olsen G. J., Davis J. J., Disz T., Robert A. Edwards R.A., Gerdes S., Parrello B., Shukla M., Vonstein V., Wattam A.R., Xia F. and Stevens R. (2014). The SEED and the rapid annotation of microbial genomes using subsystems technology (RAST). *Nucleic Acids Research*, **42(D1),** D206–D214. DOI:10.1093/nar/gkt1226
242. Pasolli E., Schiffer L., Renson A., Obenchain V., Manghi P., Truong D. T., Dowd J. B., Huttenhower C., Morgan M., Segata N. and Waldron L. (2017). Accessible, curated metagenomic data through experiment hub. *BioRxiv,* **103085,** DOI: 10.1101/103085
243. Patil K. R., Roune L. and McHardy A. C. (2012). The PhyloPythiaS web server for taxonomic assignment of metagenome sequences. *Plos One,* **7(6),** e38581. DOI:10.1371/journal.pone.0038581
244. Pearson W. R. (2013). An introduction to sequence similarity ("homology") searching. *Current Protocols in Bioinformatics,* **42,** 3.1.1-3.1.8. DOI:10.1002/0471250953.bi0301s42
245. Peck H.D. Jr. (1961). Enzymatic basis for assimilatory and dissimilatory sulfate reduction. *Journal of Bacteriology,* **82(6),** 933-9. DOI: 10.1128/jb.82.6.933-939.1961

246. Pedersen J. N., Bombar D., Paerl R. W. and Riemann L. (2018). Diazotrophs and N_2-fixation associated with particles in coastal estuarine waters. *Frontiers in Microbiology,* **9,** 2759. DOI:10.3389/fmicb.2018.02759

247. Pell J., Hintze A., Canino-Koning R., Howe A., Tiedje J. M. and Brown C. T. (2012). Scaling metagenome sequence assembly with probabilistic de Bruijn graphs. *Proceedings of the National Academy of Sciences of the United States of America,* **109(33),** 13272-13277. DOI:10.1073/pnas.1121464109

248. Peña-Ocaña B. A., Ovando-Ovando C. I., Puente-Sánchez F., Tamames J., Servín-Garcidueñas L. E., González-Toril E., Gutiérrez-Sarmiento W., Jasso-Chávez R. and Ruíz-Valdiviezo V. M. (2022). Metagenomic and metabolic analyses of poly-extreme microbiome from an active crater volcano lake. *Environmental Research,* **203,** 111862. DOI:10.1016/j.envres.2021.111862

249. Peng Y., Leung H. C., Yiu S. M. and Chin F. Y. (2012) IDBA-UD: a de novo assembler for single-cell and metagenomic sequencing data with highly uneven depth. *Bioinformatics,* **28(11),** 1420-1428. DOI:10.1093/bioinformatics/bts174

250. Pérez E. and Chebude Y. (2017). Chemical analysis of Gaet'ale, a hypersaline pond in Danakil Depression (Ethiopia): New record for the most saline water body on Earth. *Aquatic Geochemistry,* **23,** 109–117. DOI:10.1007/s10498-017-9312-z

251. Pérez C. A., Thomas F. M., Silva W. A., Aguilera R. and Armesto J. J. (2017). Biological nitrogen fixation in a post-volcanic chronosequence from south-central Chile. *Biogeochemistry,* **132,** 23-36. DOI:10.1007/s10533-016-0285-6

252. Perito B. and Cavalieri D. (2018). Innovative metagenomic approaches for detection of microbial communities involved in biodeteriorattion of cultural heritage. *IIOP*

Conference Series: Materials Science and Engineering, **364,** 01207. DOI:10.1088/1757-899x/364/1/012074

253. Plominsky A. M., Henríquez-Castillo C., Delherbe N., Podell S., Ramirez-Flandes S., Ugalde J. A., Santibañez J. F., van den Engh G., Hanselmann K., Ulloa O., De la Iglesia R., Allen E. E. and Trefault N. (2018). Distinctive archaeal composition of an artisanal crystallizer pond and functional insights into salt-saturated hypersaline environment adaptation. *Frontiers in Microbiology,* **9,** 1800. DOI: 10.3389/fmicb.2018.01800
254. Podell S., Emerson J. B., Jones C. M., Ugalde J. A., Welch S., Heidelberg K. B. and Allen E. E. (2014). Seasonal fluctuations in ionic concentrations drive microbial succession in a hypersaline lake community. *The ISME journal,* **8(5),** 979-990. DOI: 10.1038/ismej.2013.221
255. Poretsky R., Rodriguez-R L. M., Luo C., Tsementzi D. and Konstantinidis K. T. (2014). Strengths and limitations of 16S rRNA gene amplicon sequencing in revealing temporal microbial community dynamics. *Plos One,* **9(4),** e93827. DOI:10.1371/journal.pone.0093827
256. Poudel S., Colman D. R., Fixen K. R., Ledbetter R. N., Zheng Y., Pence N., Seefeldt L. C., Peters J. W., Harwood C. S. and Boyd E. S. (2018). Electron transfer to nitrogenase in different genomic and metabolic backgrounds. *Journal of Bacteriology,* **200(10),** e00757-17. DOI:10.1128/JB.00757-17
257. Quadri I., Hassani I. I., l'Haridon S., Chalopin M., Hacène H. and Jebbar M. (2016). Characterization and antimicrobial potential of extremely halophilic archaea isolated from hypersaline environments of the Algerian Sahara. *Microbiology Research,* **186(187),** 119-131. DOI:10.1016/j.micres.2016.04.003

258. Quax W. J. (2006). Bacterial enzyme. In S. Falkow, E. Rosenberg, K. Schleifer, E. Stackebrandt (Eds). The Prokaryotes: Vol. 1: Symbiotic Associations, Biotechnology, Applied Microbiology. (pp. 777-796). Springer Science and Business Media.

259. Ragsdale S. W. and Pierce E. (2008). Acetogenesis and the Wood–Ljungdahl pathway of CO_2 fixation. *Biochimica et Biophysica Acta, Proteins and Proteomics,* **1784(12),** 1873-1898. DOI:10.1016/j.bbapap.2008.08.012

260. Rampelotto P. H. (2013). Extremophiles and extreme environments. *Life,* **3(3),** 482–485. DOI:10.3390/life3030482

261. Ranjan R., Rani A., Metwally A., McGee H. S. and Perkins D. L. (2015). Analysis of the microbiome: Advantages of whole genome shotgun versus 16S amplicon sequencing. *Biochemical and Biophysical Research Communications,* **469(4),** 967-977. DOI: 10.1016/j.bbrc.2015.12.083

262. Raposo G. and Stoorvogel W. (2013). Extracellular vesicles: exosomes, microvesicles, and friends. *Journal of Cell Biology,* **200(4),** 373-83. DOI: 10.1083/jcb.201211138

263. Raza, K. (2010). Application of data mining in bioinformatics. *Indian Journal of Computer Science and Engineering,* **1(2),** 114-118. ISSN : 0976-5166

264. Reed C. J., Lewis H., Trejo E., Winston V. and Evilia C. (2013). Protein adaptations in archaeal extremophiles. *Archaea,* **373275,** 14. DOI:10.1155/2013/373275

265. Rhoads A. and Au K. F. (2015). PacBio sequencing and its applications. *Genomics, Proteomics and Bioinformatics,* **13(5),** 278-289. DOI:10.1016/j.gpb.2015.08.002

266. Rizzo J. M. and Buck M. J. (2012). Key principles and clinical applications of "next-generation" DNA sequencing. *Cancer Prevention Research,* **5(7),** 887-900. DOI: 10.1158/1940-6207

267. Rodriguez-Brito B., Rohwer F. and Edwards R. A. (2006). An application of statistics to comparative metagenomics. *BMC Bioinformatics*, **7,** 162 DOI:10.1186/1471-2105-7-162

268. Rodriguez R and Durán P. (2020). Natural holobiome engineering by using native extreme microbiome to counteract the climate change effects. *Frontiers in Bioengineering/Biotechnology,* **8,** 568. DOI:10.3389/fbioe.2020.00568

269. Rothschild L. and Mancinelli R. (2001). Life in extreme environments. *Nature,* **409,** 1092–1101. DOI:10.1038/35059215

270. Roux S., Enault F., Bronner G.L. and Debroas D. (2011). Comparison of 16S rRNA and protein-coding genes as molecular markers for assessing microbial diversity (Bacteria and Archaea) in ecosystems. *FEMS Microbiology Ecology,* **78(3),** 617–628. DOI:10.1111/j.1574-6941.2011.01190.x

271. Salwan R. and Sharma V. (2019) Proteases from extremophilic fungi: A tool for white biotechnology. In: Yadav A., Mishra S., Singh S., Gupta A. (eds) Recent Advancement in White Biotechnology Through Fungi. Fungal Biology. Springer, Cham. DOI:10.1007/978-3-030-10480-19

272. Santero E., Floriano B. and Govantes F. (2016). Harnessing the power of microbial metabolism. *Current Opinion in Microbiology,* **31,** 63–69. DOI:10.1016/j.mib.2016.03.003

273. Sanz J. L., Rodríguez N., Díaz E. E. and Amils R. (2011). Methanogenesis in the sediments of Rio Tinto, an extreme acidic river. *Environmental Microbiology,* **13(8),** 2336–2341. DOI:10.1111/j.1462-2920.2011.02504.x

274. Schiff J. A. and Fankhauser H. (1981) Assimilatory sulfate reduction. In: Bothe H., Trebst A. (eds) Biology of Inorganic Nitrogen and Sulfur. Proceedings in Life Sciences. Springer, Berlin, Heidelberg. DOI:10.1007/978-3-642-67919-3_11

275. Schmeisser C., Steele H. and Streit W. R. (2007). Metagenomics, biotechnology with non-culturable microbes. *Applied Microbiology and Biotechnology,* **75(5),** 955-962. DOI: 10.1007/s00253-007-0945-5

276. Schmidt T. M., DeLong E. F. and Pace N. R. (1991). Analysis of a marine picoplankton community by 16S rRNA gene cloning and sequencing. *Journal of Bacteriology,* **173(14),** 4371-4378. DOI:10.1128/jb.173.14.4371-4378.1991

277. Scholz C. F. P. and Kilian M. (2016). The natural history of cutaneous propionibacteria, and reclassification of selected species within the genus *Propionibacterium* to the proposed novel genera *Acidipropionibacterium gen. nov., Cutibacterium gen. nov.* and *Pseudopropionibacterium gen. nov. International Journal of Systematic and Evolutionary Microbiology,* **66(11),** 4422-4432. DOI: 10.1099/ijsem.0.001367

278. Schuchmann K. and Müller V. (2013).Direct and reversible hydrogenation of CO_2 to formate by a bacterial carbon dioxide reductase. *Science,* **342(6164),** 1382-5. DOI: 10.1126/science.1244758

279. Schuler C. G., Havig J. R. and Hamilton T. L. (2017). Hot spring microbial community composition, morphology, and carbon fixation: implications for interpreting the ancient rock record. *Frontiers in Earth Science,* **5,** 97. DOI:10.3389/feart.2017.00097

280. Sczyrba, A., Hofmann, P., Belmann, P. Koslicki D., Janssen S., Dröge J., Gregor I., Majda S., Fiedler J., Dahms E., Bremges A., Fritz A., Garrido-Oter R., Jørgensen T.

S., Shapiro N., Blood P. D., Gurevich A., Bai Y., Turaev D., DeMaere M. Z., … McHardy A. C. (2017). Critical assessment of metagenome interpretation—a benchmark of metagenomics software. *Nature Methods*, **14(11),** 1063–1071. DOI:10.1038/nmeth.4458

281. Segata N., Waldron L., Ballarini A., Narasimhan V., Jousson O. and Huttenhower C. (2012). Metagenomic microbial community profiling using unique clade-specific marker genes. *Nature methods,* **9(8),** 811-814. DOI: 10.1038/nmeth.2066

282. Segata N., Boernigen D., Tickle T .L., Morgan X. C., Garrett W. S. and Huttenhower C. (2013). Computational meta'omics for microbial community studies. *Molecular Systems Biology,* **14(9),** 666. DOI:10.1038/msb.2013.22

283. Sela-Adler M., Ronen Z., Herut B., Antler G., Vigderovich H., Eckert W. and Sivan O. (2017). Co-existence of methanogenesis and sulfate reduction with common substrates in sulfate-rich estuarine sediments. *Frontiers in Microbiology,* **8,** 766. DOI:10.3389/fmicb.2017.00766

284. Selvarajan R., Maredza A. and Tekere M. (2014). Microbial exploration in extreme conditions. In book: Metagenomic analysis and future perspectives. Chapter: 8. pages 157-181. Nova Science Publishers, Inc.Editors: Camilla Benedetti. ISBN: 978-1-63321-698-3

285. Senjarini K., Karsten U. and Schumann R. (2013). Application of fluorescence markers for the diagnosis of bacterial abundance and viability in aquatic ecosystem. *Journal of Microbiology Research,* **3(4),** 143-147. DOI:10.5923/j.microbiology.20130304.03

286. Sharpton T. J. (2014). An introduction to the analysis of shotgun metagenomic data. *Frontiers in Plant Science,* **16(5),** 209. DOI:10.3389/fpls.2014.00209

287. Sheehan K. B., Henson J. M. and Ferris M. J. (2005). *Legionella* species diversity in an acidic biofilm community in Yellowstone National Park. *Applied and Environmental Microbiology,* **71(1),** 507–511. DOI:10.1128/AEM.71.1.507-511.2005
288. Sheldon R. A. (2017). The E factor 25 years on: the rise of green chemistry and sustainability. *Green Chemistry,* **19,** 18-43. DOI:10.1039/c6gc02157c
289. Shokrollahzade S., Sharifi F., Vaseghi A., Faridounnia M. and Jahandideh S. (2015). Protein cold adaptation: Role of physico-chemical parameters in adaptation of proteins to low temperatures. *Journal of Theoretical Biology,* **383,** 130-7. DOI: 10.1016/j.jtbi.2015.07.013
290. Sieber J. R., McInerney M. J., Plugge C. M., Schink B. and Gunsalus R. P. (2010). Methanogenesis: syntrophic metabolism. In Handbook of Hydrocarbon and Lipid Microbiology; Timmis, K.N., Ed.; Springer: Berlin/Heidelberg, Germany, DOI:10.1007/978-3-540-77587-4_22
291. Silva T. P., Noyma N. P., Duque T. L. A., Gamalier J. P., Vidal L. O. Lobao L. M., Chiarini-Garcia H., Roland F. and Melo R. C. N. (2013). Visualizing aquatic bacteria by light and transmission electron microscopy. *Antonie van Leeuwenhoek,* **105(1),** 1–14. DOI: 10.1007/s10482-013-0047-6
292. Sikorski J., Tindall B.J., Lowry S., Lucas S., Nolan M., Copeland A., Glavina Del Rio T., Tice H., Cheng J. F., Han C., Pitluck S., Liolios K., Ivanova N., Mavromatis K., Mikhailova N., Pati A., Goodwin L., Chen A., Palaniapan K., Land M., … Lapidus A. (2010). Complete genome sequence of *Meiothermus silvanus* type strain (VI-R2T). *Standards in Genomic Sciences,* **3(1),** 37–46. DOI:10.4056/sigs.1042812

293. Simachew A., Lanzén A., Gessesse A., and Øvreås L. (2016). Prokaryotic community diversity along an increasing salt gradient in a soda ash concentration pond. *Microbial Ecology,* **71(2),** 326-338. DOI:10.1007/s00248-015-0675-7

294. Slonczewski J. L., Gonzalez T. N., Bartholomew F. M. and Holt N. J. (1987). Mu d-directed lacZ fusions regulated by low pH in *Escherichia coli*. *Journal of bacteriology,* **169(7),** 3001–3006. DOI:10.1128/jb.169.7.3001-3006.1987

295. Smith A. A. and Plazas M. C. (2011). Insilico characterization and Homology Modeling of Cyanobacterial Phosphoenolpyruvate Carboxylase Enzymes with computiational Tools and Bioinformatic Servers. *American Journal of Biochemistry and Molecular Biology.* **1(4):** 319-336.

296. Song Z. Q., Wang F. P., Zhi X. Y., Chen J. Q., Zhou E. M., Liang F., Xiao X., Tang S. K., Jiang H. C., Zhang C. L., Dong H. and Li W. J. (2013). Bacterial and archaeal diversities in Yunnan and Tibetan hot springs, China. *Environmental Microbiology,* **15(4),** 1160-75. DOI: 10.1111/1462-2920.12025

297. Sonia J., Dorra A., Hajer Ben H. and Samir B. (2016). Biocatalysts: application and engineering for industrial purposes. *Critical Reviews in Biotechnology,* **36(2),** 246-258. DOI: 10.3109/07388551.2014.950550

298. Sorokin D. Y., Berben T., Melton E. D., Overmars L., Vavourakis C. D., and Muyzer G. (2014). Microbial diversity and biogeochemical cycling in soda lakes. *Extremophiles,* **18(5),** 791-809. DOI:10.1007/s00792-014-0670-9

299. Sorokin D. Y., Makarova K. S., Abbas B., Ferrer M., Golyshin P. N., Galinski E. A., Ciordia S., Mena M. C., Merkel A. Y., Wolf Y. I., van Loosdrecht M. and Koonin E. V. (2017). Discovery of extremely halophilic, methyl-reducing euryarchaea provides

insights into the evolutionary origin of methanogenesis. *Nature Microbiology,* **2,** 17081. DOI: 10.1038/nmicrobiol.2017.81.

300. Spear J. R., Walker J. J. and Pace N. R (2006). Microbial ecology and energetics in Yellowstone hot springs. *Yellowstone Science,* **14(1).**
301. Stetter K. O. (1982). Ultrathin mycelia-forming organisms from submarine volcanic areas having an optimum growth temperature of 105 °C. *Nature,* **300,** 258–260. DOI:10.1038/300258a0
302. Stivaletta N., Barbieri R. and Billi D. (2012). Microbial colonization of the salt deposits in the driest place of the tacama Desert (Chile). *Origins of Life and Evolution of the Biosphere*, **42,** 187-200. DOI:10.1007/s11084-012-9289-y
303. Stout L. M., Blake R. E., Greenwood J. P., Martini A. M. and Rose E. C. (2009). Microbial diversity of boron-rich volcanic hot springs of St. Lucia, Lesser Antilles. *FEMS Microbiology Ecology,* **70(3),** 402–412. DOI:10.1111/j.1574-6941.2009.00780.x
304. Su C., Lei L., Duan Y., Zhang K. and Yang J. (2011). Culture-independent methods for studying environmental microorganisms: Methods, application, and perspective. *Applied Microbiology and Biotechnology,* **93,** 993-1003. DOI:10.1007/s00253-011-3800-7
305. Tan B., Jane F. S., Laban N. A., Dong X., Sensen C. W., Foght J. and Gieg L. M. (2015). Comparative analysis of metagenomes from three methanogenic hydrocarbon-degrading enrichment cultures with 41 environmental samples. *The* I*SME Journal,* **9(9),** 2028-2045. DOI: 10.1038/ismej.2015.22
306. Tan K. Y., Dutta A., Tan T. K., Hari R., Othman R. and Choo S. W. (2020). Comprehensive genome analysis of a pangolin-associated *Paraburkholderia*

fungorum provides new insights into its secretion systems and virulence. *PeerJ,* **8,** e9733. DOI: 10.7717/peerj.9733

307. Teo Y. V., and Neretti N. (2016). A comparative study of metagenomics analysis pipelines at the species level. *BioRxiv.* **081141**. DOI: 10.1101/081141
308. Thiel V. (2011) Extreme environments. In: Reitner J., Thiel V. (eds) Encyclopedia of Geobiology. Encyclopedia of Earth Sciences Series. Springer, Dordrecht. DOI:10.1007/978-1-4020-9212-1_87
309. Thomas T., Gilbert J. and Meyer F. (2012). Metagenomics - a guide from sampling to data analysis. *Microbial Informatics and Experimentation,* **2(1)**: 3. DOI: 10.1186/2042-5783-2-3
310. Tikh I. and Schmidt-Dannert C. (2013). Towards engineered light-energy conversion in non photosynthetic microorganisms. In Synthetic Biology (pp. 303-316). Elsevier Inc.. DOI: 10.1016/B978-0-12-394430-6.00016-9
311. Tilahun L., Asfawossen A., Wessel G. M. and Simachew A. (2021^a). In silico determination of nitrogen metabolism in microbes from extreme conditions using metagenomics. *Archives of Microbiology*. **203(5),** 2521–2540. DOI: 10.1007/s00203-021-02227-4
312. Tilahun L., Asfawossen A., Wessel G. M. and Simachew A. (2021^b). Prediction of genes that function in methanogenesis and CO_2 Pathways in Extremophiles. *Microorganisms,* **9,** 2211. DOI:10.3390/microorganisms9112211
313. Toyofuku M., Nomura N. and Eberl L. (2019). Types and origins of bacterial membrane vesicles. *Nature Review in Microbiology,* **17,** 13–24. DOI:10.1038/s41579-018-0112-2

314. Tran Q. and Phan V. (2020). Assembling reads improves taxonomic classification of species. *Genes,* **11,** 946. DOI:10.3390/genes11080946

315. Truong D. T., Franzosa E. A., Tickle T. L., Scholz M., Weingart G., Pasolli E., Tett A., Huttenhower C. and Segata N. (2015). MetaPhlAn2 for enhanced metagenomic taxonomic profiling. *Nature Methods,* **12(10),** 902-903. DOI: 10.1038/nmeth.3589

316. Tsai S. J. and Yeates T. O. (2011). Bacterial microcompartments insights into the structure, mechanism, and engineering applications. *Progress in Molecular Biology and Translational Science,* **103,** 1-20. DOI: 10.1016/B978-0-12-415906-8.00008-X

317. Turturici G., Tinnirello R., Sconzo G. and Geraci F. (2014). Extracellular membrane vesicles as a mechanism of cell-to-cell communication: advantages and disadvantages. *American journal of physiology. Cell physiology,* **306(7),** C621-33. DOI: 10.1152/ajpcell.00228.2013

318. Tyler A. D., Matasej L., Urfano C. J. Schmidt L., Antonation K. S., Mulvey M. R. and Corbett C. R. (2018). Evaluation of Oxford Nanopore's MinION sequencing device for microbial whole genome sequencing applications. *Scientific Reports,* **8,** 10931. DOI:10.1038/s41598-018-29334-5

319. Uribe-Lorío L., Brenes-Guillén L., Hernández-Ascencio W., Mora-Amador R., González G., Ramírez-Umaña C.J., Díez B. and Pedrós-Alió C. (2019). The influence of temperature and pH on bacterial community composition of microbial mats in hot springs from Costa Rica. *Microbiology Open*. **8(10),** e893. DOI:10.1002/mbo3.893

320. Usman K, Shehu I., Dogara A. M. and Idris Z. L. (2017). In-silico characterization and mutagenesis of β-Glucosidase from *Anoxybacillus sp. SK3-4*. *American Journal of Bioscience and Bioengineering,* **5(1),** 1-7. DOI:10.11648/j.bio.20170501.11

321. van den Burg B. (2003). Extremophiles as a source for novel enzymes. *Current Opinion in Microbiology,* **6(3),** 213-8. DOI:10.1016/s1369-5274(03)00060-2

322. van Elsas J. D., Semenov A. V., Costa R. and Trevors J. T. (2011). Survival of *Escherichia coli* in the environment: fundamental and public health aspects. *The ISME journal,* **5(2),** 173–183. DOI:0.1038/ismej.2010.80

323. Van Kranendonk M. J., Deamer D. W., and Djokic T. (2017) Life springs. Springs. *Scientific American,* **317(2),** 28–35. DOI:10.1038/scientificamerican0817-28

324. Vanwonterghem I., Evans P., Parks D., Jensen P.D., Woodcroft B. J., Hugenholtz P. and Tyson G. W. (2016). Methylotrophic methanogenesis discovered in the archaeal phylum Verstraetearcheaeota. *Nature Microbiology,* **1,** 16170. DOI: 10.1038/nmicrobiol.2016.170

325. Ventosa A., and Arahal D. (2009). Physico-chemical characteristics of hypersaline environments and their biodiversity. In Extremophiles Volume 2 (pp 245-251). Encyclopedia of Life Support System. UNESCO-EOLSS.

326. Ventosa A., Gutiérrez M. C., Kamekura M., Zvyagintseva I. S., and Oren A. (2004). Taxonomic study of *Halorubrum distributum* and proposal of *Halorubrum terrestre sp. nov. International Journal of Systematic and Evolutionary Microbiology,* **54(2),** 389-392. DOI:10.1099/ijs.0.02621-0

327. Vereb V., van Wyk de Vries B., Hagos M. and Karátson D. (2020). Geoheritage and Resilience of Dallol and the Northern Danakil Depression in Ethiopia. *Geoheritage,* **12,** 82. DOI:10.1007/s12371-020-00499-8

328. Vio S. A., García S. S., Casajus V., Arango J. S., Galar M. L., Bernabeu P. R. and Luna M. F. (2020). *Paraburkholderia*: beneficial microbes in agro-ecology. *Academic Press*, **15,** 271-311. DOI:10.1016/B978-0-12-823414-3.00015-0

329. Viver T., Orellana L., González-Torres P., Díaz S., Urdiain M., Farías M. E., Benes V., Kaempfer P., Shahinpei A., Amoozegar M. A., Amann R., Josefa Antón J., Konstantinidis K. T. and Rosselló-Móra R. (2018). Genomic comparison between members of the Salinibacteraceae family, and description of a new species of *Salinibacter* (*Salinibacter altiplanensis sp. nov.*) isolated from high altitude hypersaline environments of the Argentinian Altiplano. *Systematic and Applied Microbiology,* **41(3),** 198-212. DOI:10.1016/j.syapm.2017.12.004

330. Wächtershäuser G. (2006). From volcanic origins of chemoautotrophic life to Bacteria, Archaea and Eukarya. *Philosophical Transactions of the Royal Society of London. Series B, Biological sciences,* **361(1474),** 1787–1808. DOI: 10.1098/rstb.2006.1904

331. Wang P., Song P., Jin M. and Zhu G. (2013). Isocitrate dehydrogenase from *Streptococcus mutans:* biochemical properties and evaluation of a putative phosphorylation site at Ser102. *Plos One*. **8(3),** e58918. DOI: 10.1371/journal.pone.0058918

332. Wang Q., Cen Z. and Zhao J. (2015). The survival mechanisms of thermophiles at high temperatures: an angle of Omics. *Physiology,* **30(2),** 97-106. DOI:10.1152/physiol.00066.2013

333. Wang Y., Wegener G., Williams T. A., Xie R., Hou J., Wang X. and Xiao X. (2021). A methylotrophic origin of methanogenesis and early divergence of anaerobic multicarbon alkane metabolism. *Science Advances,* **7(7),** eabd7180. DOI: 10.1126/sciadv.abd7180

334. Waterhouse A., Bertoni M., Bienert S., Studer G., Tauriello G., Gumienny R., Heer F. T., de Beer T. A. P., Rempfer C., Bordoli L., Lepore R. and Schwede T. (2018).

Swiss-Model: homology modelling of protein structures and complexes. *Nucleic Acids Research,* **46,** W296-W303. DOI:10.1093/nar/gky427

335. Watson N. Dunyak D., Rosey E., Sloncz E. and Olson A. (1992). Identification of elements involved in transcriptional regulation of the *Escherichia coli* cad operon by external pH. *Journal of Bacteriology,* **174(2**), 530-540. DOI: 0021-9193/92/020530-11$02.00/0
336. Weber C. F. and King G. M. (2017).Volcanic soils as sources of novel CO-oxidizing *Paraburkholderia* and *Burkholderia: Paraburkholderia hiiakae* sp. nov., *Paraburkholderia metrosideri* sp. nov., *Paraburkholderia paradisi* sp. nov., *Paraburkholderia peleae* sp. nov., and *Burkholderia alpina* sp. nov. a member of the *Burkholderia cepacia* complex. *Frontiers in Microbiology,* **8,** 207. DOI:10.3389/fmicb.2017.00207
337. Whitman W. B., Coleman D. C. and Wiebe W. J. (1998). Prokaryotes: the unseen majority. *Proceedings of the National Academy of Sciences of the United States of America,* **95(12),** 6578–6583. DOI: 10.1073/pnas.95.12.6578
338. Widdison P. E. and Burt T. P. (2008). Nitrogen Cycle. Editor(s): Sven Erik Jørgensen, Brian D. Fath. Encyclopedia of Ecology, Academic Press. 2526-2533. ISBN 9780080454054. DOI:10.1016/B978-008045405-4.00750-3
339. Wiklander O. P. B., Brennan M. A., Lotvall J., Breakefield X. O. and Andaloussi S. E. (2019). Advances in therapeutic applications of extracellular vesicles. *Science Translational Medicine,* **11(492),** eaav8521. DOI:10.1126/scitranslmed.aav8521
340. Wilkins L. G. E., Ettinger C. L., Jospin G. and Eisenet J. A. (2019). Metagenome assembled genomes provide new insight into the microbial diversity of two thermal

pools in Kamchatka, Russia. *Scientific Reports,* **9,** 3059. DOI:10.1038/s41598-019-39576-6

341. Woldekiros H. S. (2019). The route most traveled: The Afar salt trail. north Ethiopia. *Páginas,* **51(1),** 95-110. DOI:10.4067/S0717-73562019005000502
342. Yang J., Ma L., Jiang H. Wu G. and Dong H. (2016). Salinity shapes microbial diversity and community structure in surface sediments of the Qinghai-Tibetan Lakes. *Scientific Reports,* **6,** 25078. DOI:10.1038/srep25078
343. Yang Y., Cui H. L., Zhou P. J., and Liu S. J. (2007). *Haloarcula amylolytica* sp. nov., an extremely halophilic archaeon isolated from Aibi salt lake in Xin-Jiang, China. *International Journal of Systematic and Evolutionary Microbiology,* **57(1),**103-106. DOI:10.1099/ijs.0.64647-0
344. Yeh H. and Lee J. (2018). Stable hydrogen and oxygen isotopes for groundwater sources of Penghu Islands, Taiwan. *Geosciences,* **8,** 84. DOI:10.3390/geosciences8030084
345. Yokoyama F., Kawamoto J., Imai T. and Tatsuo K. (2017). Characterization of extracellular membrane vesicles of an Antarctic bacterium, S*hewanella livingstonensis Ac10,* and their enhanced production by alteration of phospholipid composition. *Extremophiles,* **21,** 723–731. DOI:10.1007/s00792-017-0937-z
346. Yuan H., Yuan J., You Y., Zhang B., Wu Y., Huang S. and Zhang Y. (2021). Simultaneous ammonium and sulfate biotransformation driven by aeration: Nitrogen/sulfur metabolism and metagenome-based microbial ecology. *Science of the Total Environment,* **794,**148650. DOI:10.1016/j.scitotenv.2021.148650
347. Yue, Y., Huang, H., Qi, Z. Dou H., Liu X., Han T., Chen Y., Song X., Zhang Y. and Tu J. (2020). Evaluating metagenomics tools for genome binning with real

metagenomic datasets and CAMI datasets. *BMC Bioinformatics,* **21(1),** 334. DOI:10.1186/s12859-020-03667-3

348. Zheng L., White R. H., Cash V. L., Jack R. F. and Dean D. R. (1993). Cysteine desulfurase activity indicates a role for NifS in metallocluster biosynthesis. *Proceedings of the National Academy of Sciences of the United States of America,* **90 (7),** 2754-2758. DOI: 10.1073/pnas.90.7.2754
349. Zhou J., Bruns M. A. and Tiedje J. M. (1996). DNA recovery from soils of diverse composition. *Applied and Environmental Microbiology,* **62(2),** 316–322. DOI: 10.1128/aem.62.2.316-322.1996

Appendix 1: Optimization of DNA extraction protocols for Dallol sample

For Dallol sample, the 1% CTAB-SDS DNA extraction method from Mitchell and Takacs-Vesbach, 2008 was modified and used for optimum eDNA extraction. The 1% CTAB-SDS DNA extraction method from Mitchell and Takacs-Vesbach, 2008 was proven to be ideal for extracting eDNA from acidic hot springs at Yellowstone National Park, USA. Nonetheless for DAL sample, immediate oxidation occurs upon addition of equal amount of phenol:chloroform and the color of the solution was observed to change into deep purple red. In order to check whether the method adapted from Mitchell and Takacs-Vesbach, 2008 was optimally lysing the cells to release genomic materials, 1ml of 'Dallol sample containing 1% CTAB-SDS buffer' was taken at the end of the last incubation period and was directly stained using FM™ 1-43 Dye (N-(3-Triethylammoniumpropyl)-4-(4-(Dibutylamino) Styryl) Pyridinium Dibromide), Invitrogen™ and Hoechst 33342 Solution, Thermo Scientific™, in accordance to the manufacturers instructions. According to the catalog supplied by the manufacturer, Hoechst 33342 is a cell-permeable DNA stain used for specifically staining the nuclei of living or fixed cells, and that is excited by ultraviolet light and emits blue fluorescence at 460 to 490 nm. On the other hand, FM™ 1-43 Dye, Invitrogen™ is known to insert into the outer leaflet of the cell membrane and small secretory vesicles where it becomes intensely fluorescent. Cellular imaging was done using Zeiss Axioplan Fluorescence Microscope with Hamamatsu orca ER optical camera C4742-95 at PrIMO lab.

Appendix 2: MetaphlAn 2 taxonomy abundance profile table

ID	LTW0001_out	LTW0002_out	LTW0003_out	LTW0004_out	LTW0005_out	LTW0006_out	LTW0007_out
k__Archaea	0.04	0.08	0	94.06	95.3	97.27	95.11
k__Archaea|p__Euryarchaeota	0.04	0.08	0	94.06	95.3	97.27	95.11
k__Archaea|p__Euryarchaeota|c__Halobacteria	0.04	0.08	0	94.06	95.3	97.27	95.11
k__Archaea|p__Euryarchaeota|c__Halobacteria|o__Halobacteriales	0.04	0.05	0	14.49	32.92	11.7	74.56
k__Archaea|p__Euryarchaeota|c__Halobacteria|o__Halobacteriales|f__Haloarculaceae	0	0	0	2.13	3.05	0.98	3.11
k__Archaea|p__Euryarchaeota|c__Halobacteria|o__Halobacteriales|f__Haloarculaceae|g__Haloarcula	0	0	0	1.98	2.17	0.93	1.41
k__Archaea|p__Euryarchaeota|c__Halobacteria|o__Halobacteriales|f__Haloarculaceae|g__Haloarcula|s__Haloarcula_amylolytica	0	0	0	0.1	0.05	0.05	0.03
k__Archaea|p__Euryarchaeota|c__Halobacteria|o__Halobacteriales|f__Haloarculaceae|g__Haloarcula|s__Haloarcula_amylolytica|t__GCF_000336615	0	0	0	0.1	0.05	0.05	0.03
k__Archaea|p__Euryarchaeota|c__Halobacteria|o__Halobacteriales|f__Haloarculaceae|g__Haloarcula|s__Haloarcula_californiae	0	0	0	0.18	0.35	0.1	0.02
k__Archaea|p__Euryarchaeota|c__Halobacteria|o__Halobacteriales|f__Haloarculaceae|g__Haloarcula|s__Haloarcula_californiae|t__GCF_000337755	0	0	0	0.18	0.35	0.1	0.02
k__Archaea|p__Euryarchaeota|c__Halobacteria|o__Halobacteriales|f__Haloarculaceae|g__Haloarcula|s__Haloarcula_hispanica	0	0	0	0.11	0.16	0.04	0.16
k__Archaea|p__Euryarchaeota|c__Halobacteria|o__Halobacteriales|f__Haloarculaceae|g__Haloarcula|s__Haloarcula_hispanica|t__GCF_000223905	0	0	0	0.11	0.16	0.04	0.16
k__Archaea|p__Euryarchaeota|c__Halobacteria|o__Halobacteriales|f__Haloarculaceae|g__Haloarcula|s__Haloarcula_japonica	0	0	0	1.32	1.34	0.58	1.13
k__Archaea|p__Euryarchaeota|c__Halobacteria|o__Halobacteriales|f__Haloarculaceae|g__Haloarcula|s__Haloarcula_japonica|t__GCF_000336635	0	0	0	1.32	1.34	0.58	1.13
k__Archaea|p__Euryarchaeota|c__Halobacteria|o__Halobacteriales|f__Haloarculaceae|g__Haloarcula|	0	0	0	0.12	0.17	0.09	0.02

s__Haloarcula_marismortui							
k__Archaea|p__Euryarchaeota|c__Halobacteria|o__Halobacteriales|f__Haloarculaceae|g__Haloarcula|s__Haloarcula_marismortui|t__GCF_000011085	0	0	0	0.12	0.17	0.09	0.02
k__Archaea|p__Euryarchaeota|c__Halobacteria|o__Halobacteriales|f__Haloarculaceae|g__Haloarcula|s__Haloarcula_sinaiiensis	0	0	0	0.06	0.04	0.04	0.01
k__Archaea|p__Euryarchaeota|c__Halobacteria|o__Halobacteriales|f__Haloarculaceae|g__Haloarcula|s__Haloarcula_sinaiiensis|t__GCF_000337275	0	0	0	0.06	0.04	0.04	0.01
k__Archaea|p__Euryarchaeota|c__Halobacteria|o__Halobacteriales|f__Haloarculaceae|g__Haloarcula|s__Haloarcula_vallismortis	0	0	0	0.09	0.07	0.03	0.03
k__Archaea|p__Euryarchaeota|c__Halobacteria|o__Halobacteriales|f__Haloarculaceae|g__Haloarcula|s__Haloarcula_vallismortis|t__GCF_000337775	0	0	0	0.09	0.07	0.03	0.03
k__Archaea|p__Euryarchaeota|c__Halobacteria|o__Halobacteriales|f__Haloarculaceae|g__Halorhabdus	0	0	0	0.11	0.4	0.04	1.68
k__Archaea|p__Euryarchaeota|c__Halobacteria|o__Halobacteriales|f__Haloarculaceae|g__Halorhabdus|s__Halorhabdus_tiamatea	0	0	0	0	0.04	0	0.16
k__Archaea|p__Euryarchaeota|c__Halobacteria|o__Halobacteriales|f__Haloarculaceae|g__Halorhabdus|s__Halorhabdus_tiamatea|t__GCF_000215915	0	0	0	0	0.04	0	0.16
k__Archaea|p__Euryarchaeota|c__Halobacteria|o__Halobacteriales|f__Haloarculaceae|g__Halorhabdus|s__Halorhabdus_unclassified	0	0	0	0.11	0.36	0.04	1.46
k__Archaea|p__Euryarchaeota|c__Halobacteria|o__Halobacteriales|f__Haloarculaceae|g__Halorhabdus|s__Halorhabdus_utahensis	0	0	0	0	0	0	0.06
k__Archaea|p__Euryarchaeota|c__Halobacteria|o__Halobacteriales|f__Haloarculaceae|g__Halorhabdus|s__Halorhabdus_utahensis|t__GCF_000023945	0	0	0	0	0	0	0.06
k__Archaea|p__Euryarchaeota|c__Halobacteria|o__Halobacteriales|f__Haloarculaceae|g__Halosimplex	0	0	0	0.03	0.07	0.01	0.01
k__Archaea|p__Euryarchaeota|c__Halobacteria|o__Halobacteriales|f__Haloarculaceae|g__Halosimplex|s__Halosimplex_carlsbadense	0	0	0	0.03	0.07	0.01	0.01
k__Archaea|p__Euryarchaeota|c__Halobacteria|o__Halobacteriales|f__Haloarculaceae|g__Halosimplex|s__Halosimplex_carlsbadense|t__GCF_000337455	0	0	0	0.03	0.07	0.01	0.01
k__Archaea|p__Euryarchaeota|c__Halobacteria|o__Halobacteriales|f__Halobacteriaceae	0.04	0.03	0	7.34	24.16	6.64	68.21
k__Archaea|p__Euryarchaeota|c__Halobacteria|o__Halobacteriales|f__Halobacteriaceae|g__Halarchaeum	0.04	0	0	0	0.07	0	17.46

<table>
<tr><td>k__Archaea|p__Euryarchaeota|c__Halobacteria|o__Halobacteriales|f__Halobacteriaceae|g__Halarchaeum|s__Halarchaeum_acidiphilum</td><td>0.04</td><td>0</td><td>0</td><td>0</td><td>0.07</td><td>0</td><td>17.46</td></tr>
<tr><td>k__Archaea|p__Euryarchaeota|c__Halobacteria|o__Halobacteriales|f__Halobacteriaceae|g__Halarchaeum|s__Halarchaeum_acidiphilum|t__Halarchaeum_acidiphilum_unclassified</td><td>0.04</td><td>0</td><td>0</td><td>0</td><td>0.07</td><td>0</td><td>17.46</td></tr>
<tr><td>k__Archaea|p__Euryarchaeota|c__Halobacteria|o__Halobacteriales|f__Halobacteriaceae|g__Halobacterium</td><td>0</td><td>0.03</td><td>0</td><td>7.31</td><td>24.08</td><td>6.63</td><td>50.74</td></tr>
<tr><td>k__Archaea|p__Euryarchaeota|c__Halobacteria|o__Halobacteriales|f__Halobacteriaceae|g__Halobacterium|s__Halobacterium_unclassified</td><td>0</td><td>0.03</td><td>0</td><td>7.31</td><td>24.08</td><td>6.63</td><td>50.74</td></tr>
<tr><td>k__Archaea|p__Euryarchaeota|c__Halobacteria|o__Halobacteriales|f__Halobacteriaceae|g__Halomicrobium</td><td>0</td><td>0</td><td>0</td><td>0.02</td><td>0.01</td><td>0.01</td><td>0.01</td></tr>
<tr><td>k__Archaea|p__Euryarchaeota|c__Halobacteria|o__Halobacteriales|f__Halobacteriaceae|g__Halomicrobium|s__Halomicrobium_katesii</td><td>0</td><td>0</td><td>0</td><td>0.02</td><td>0.01</td><td>0.01</td><td>0.01</td></tr>
<tr><td>k__Archaea|p__Euryarchaeota|c__Halobacteria|o__Halobacteriales|f__Halobacteriaceae|g__Halomicrobium|s__Halomicrobium_katesii|t__GCF_000379085</td><td>0</td><td>0</td><td>0</td><td>0.02</td><td>0.01</td><td>0.01</td><td>0.01</td></tr>
<tr><td>k__Archaea|p__Euryarchaeota|c__Halobacteria|o__Halobacteriales|f__Halobacteriaceae|g__Halomicrobium|s__Halomicrobium_unclassified</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td></tr>
<tr><td>k__Archaea|p__Euryarchaeota|c__Halobacteria|o__Halobacteriales|f__Halococcaceae</td><td>0</td><td>0.02</td><td>0</td><td>5.03</td><td>5.72</td><td>4.07</td><td>3.24</td></tr>
<tr><td>k__Archaea|p__Euryarchaeota|c__Halobacteria|o__Halobacteriales|f__Halococcaceae|g__Halococcus</td><td>0</td><td>0.02</td><td>0</td><td>5.03</td><td>5.72</td><td>4.07</td><td>3.24</td></tr>
<tr><td>k__Archaea|p__Euryarchaeota|c__Halobacteria|o__Halobacteriales|f__Halococcaceae|g__Halococcus|s__Halococcus_unclassified</td><td>0</td><td>0.02</td><td>0</td><td>5.03</td><td>5.72</td><td>4.07</td><td>3.24</td></tr>
<tr><td>k__Archaea|p__Euryarchaeota|c__Halobacteria|o__Haloferacales</td><td>0</td><td>0.03</td><td>0</td><td>78.97</td><td>61.65</td><td>84.73</td><td>20.06</td></tr>
<tr><td>k__Archaea|p__Euryarchaeota|c__Halobacteria|o__Haloferacales|f__Haloferacaceae</td><td>0</td><td>0</td><td>0</td><td>0.8</td><td>0.81</td><td>0.45</td><td>0.51</td></tr>
<tr><td>k__Archaea|p__Euryarchaeota|c__Halobacteria|o__Haloferacales|f__Haloferacaceae|g__Haloferax</td><td>0</td><td>0</td><td>0</td><td>0.71</td><td>0.77</td><td>0.43</td><td>0.5</td></tr>
<tr><td>k__Archaea|p__Euryarchaeota|c__Halobacteria|o__Haloferacales|f__Haloferacaceae|g__Haloferax|s__Haloferax_alexandrinus</td><td>0</td><td>0</td><td>0</td><td>0.04</td><td>0.06</td><td>0.04</td><td>0.12</td></tr>
<tr><td>k__Archaea|p__Euryarchaeota|c__Halobacteria|o__Haloferacales|f__Haloferacaceae|g__Haloferax|s__Haloferax_alexandrinus|t__GCF_000336735</td><td>0</td><td>0</td><td>0</td><td>0.04</td><td>0.06</td><td>0.04</td><td>0.12</td></tr>
<tr><td>k__Archaea|p__Euryarchaeota|c__Halobacteria|o__Haloferacales|f__Haloferacaceae|g__Haloferax|s__Haloferax_denitrificans</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0.02</td><td>0.01</td><td>0.01</td></tr>
<tr><td>**k__Archaea|p__Euryarchaeota|c__Halobacteria|o__Haloferacales|f__Haloferacaceae|g__Haloferax|**</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0.02</td><td>0.01</td><td>0.01</td></tr>
</table>

s__Haloferax_denitrificans|t__GCF_000337795							
k__Archaea|p__Euryarchaeota|c__Halobacteria|o__Haloferacales|f__Haloferacaceae|g__Haloferax|s__Haloferax_larsenii	0	0	0	0.55	0.32	0.3	0.22
k__Archaea|p__Euryarchaeota|c__Halobacteria|o__Haloferacales|f__Haloferacaceae|g__Haloferax|s__Haloferax_larsenii| t__GCF_000336955	0	0	0	0.55	0.32	0.3	0.22
k__Archaea|p__Euryarchaeota|c__Halobacteria|o__Haloferacales|f__Haloferacaceae|g__Haloferax|s__Haloferax_lucentense	0	0	0	0.03	0.06	0.02	0.03
k__Archaea|p__Euryarchaeota|c__Halobacteria|o__Haloferacales|f__Haloferacaceae|g__Haloferax|s__Haloferax_lucentense| t__GCF_000336795	0	0	0	0.03	0.06	0.02	0.03
k__Archaea|p__Euryarchaeota|c__Halobacteria|o__Haloferacales|f__Haloferacaceae|g__Haloferax|s__Haloferax_prahovense	0	0	0	0	0.01	0	0.02
k__Archaea|p__Euryarchaeota|c__Halobacteria|o__Haloferacales|f__Haloferacaceae|g__Haloferax|s__Haloferax_prahovense| t__GCF_000336815	0	0	0	0	0.01	0	0.02
k__Archaea|p__Euryarchaeota|c__Halobacteria|o__Haloferacales|f__Haloferacaceae|g__Haloferax|s__Haloferax_sulfurifontis	0	0	0	0.05	0.07	0.03	0.04
k__Archaea|p__Euryarchaeota|c__Halobacteria|o__Haloferacales|f__Haloferacaceae|g__Haloferax|s__Haloferax_sulfurifontis| t__GCF_000337835	0	0	0	0.05	0.07	0.03	0.04
k__Archaea|p__Euryarchaeota|c__Halobacteria|o__Haloferacales|f__Haloferacaceae|g__Haloferax|s__Haloferax_unclassified	0	0	0	0	0.15	0	0
k__Archaea|p__Euryarchaeota|c__Halobacteria|o__Haloferacales|f__Haloferacaceae|g__Haloferax|s__Haloferax_volcanii	0	0	0	0.03	0.08	0.04	0.08
k__Archaea|p__Euryarchaeota|c__Halobacteria|o__Haloferacales|f__Haloferacaceaeg__Haloferax|s__Haloferax_volcanii| t__Haloferax_volcanii_unclassified	0	0	0	0.03	0.08	0.04	0.08
k__Archaea|p__Euryarchaeota|c__Halobacteria|o__Haloferacales|f__Haloferacaceae|g__Haloquadratum	0	0	0	0.09	0.04	0.02	0.01
k__Archaea|p__Euryarchaeota|c__Halobacteria|o__Haloferacales|f__Haloferacaceae|g__Haloquadratum| s__Haloquadratum_unclassified	0	0	0	0.09	0.04	0.02	0.01
k__Archaea|p__Euryarchaeota|c__Halobacteria|o__Haloferacales|f__Halorubraceae	0	0.03	0	78.16	60.84	84.28	19.55
k__Archaea|p__Euryarchaeota|c__Halobacteria|o__Haloferacales|f__Halorubraceae|g__Halonotius	0	0.03	0	76.87	58.74	82.96	18.2
k__Archaea|p__Euryarchaeota|c__Halobacteria|o__Haloferacales|f__Halorubraceae|g__Halonotius|s__Halonotius_sp_J07HN6	0	0	0	76.87	58.74	79.61	18.2
k__Archaea|p__Euryarchaeota|c__Halobacteria|o__Haloferacales|f__Halorubraceae|g__Halonotius|	0	0	0	76.87	58.74	79.61	18.2

s__Halonotius_sp_J07HN6|t__GCF_000416025							
k__Archaea|p__Euryarchaeota|c__Halobacteria|o__Haloferacales|f__Halorubraceae|g__Halonotius|s__Halonotius_unclassified	0	0.03	0	0	0	3.35	0
k__Archaea|p__Euryarchaeota|c__Halobacteria|o__Haloferacales|f__Halorubraceae|g__Halorubrum	0	0	0	1.3	2.1	1.32	1.35
k__Archaea|p__Euryarchaeota|c__Halobacteria|o__Haloferacales|f__Halorubraceae|g__Halorubrum|s__Halorubrum_distributum	0	0	0	0.01	0.01	0	0
k__Archaea|p__Euryarchaeota|c__Halobacteria|o__Haloferacales|f__Halorubraceae|g__Halorubrum|s__Halorubrum_distributum|t__Halorubrum_distributum_unclassified	0	0	0	0.01	0.01	0	0
k__Archaea|p__Euryarchaeota|c__Halobacteria|o__Haloferacales|f__Halorubraceae|g__Halorubrum|s__Halorubrum_hochstenium	0	0	0	0.28	0.17	0.13	0.1
k__Archaea|p__Euryarchaeota|c__Halobacteria|o__Haloferacales|f__Halorubraceae|g__Halorubrum|s__Halorubrum_hochstenium|t__GCF_000337075	0	0	0	0.28	0.17	0.13	0.1
k__Archaea|p__Euryarchaeota|c__Halobacteria|o__Haloferacales|f__Halorubraceae|g__Halorubrum|s__Halorubrum_litoreum	0	0	0	0.06	0.01	0.03	0.01
k__Archaea|p__Euryarchaeota|c__Halobacteria|o__Haloferacales|f__Halorubraceae|g__Halorubrum|s__Halorubrum_litoreum|t__GCF_000337395	0	0	0	0.06	0.01	0.03	0.01
k__Archaea|p__Euryarchaeota|c__Halobacteria|o__Haloferacales|f__Halorubraceae|g__Halorubrum|s__Halorubrum_terrestre	0	0	0	0.01	0	0.01	0
k__Archaea|p__Euryarchaeota|c__Halobacteria|o__Haloferacales|f__Halorubraceae|g__Halorubrum|s__Halorubrum_terrestre|t__GCF_000337435	0	0	0	0.01	0	0.01	0
k__Archaea|p__Euryarchaeota|c__Halobacteria|o__Haloferacales|f__Halorubraceae|g__Halorubrum|s__Halorubrum_unclassified	0	0	0	0.94	1.91	1.14	1.24
k__Archaea|p__Euryarchaeota|c__Halobacteria|o__Natrialbales	0	0	0	0.6	0.72	0.83	0.49
k__Archaea|p__Euryarchaeota|c__Halobacteria|o__Natrialbales|f__Natrialbaceae	0	0	0	0.6	0.72	0.83	0.49
k__Archaea|p__Euryarchaeota|c__Halobacteria|o__Natrialbales|f__Natrialbaceae|g__Halobiforma	0	0	0	0.26	0.29	0.26	0.23
k__Archaea|p__Euryarchaeota|c__Halobacteria|o__Natrialbales|f__Natrialbaceae|g__Halobiforma|s__Halobiforma_unclassified	0	0	0	0.26	0.29	0.26	0.23
k__Archaea|p__Euryarchaeota|c__Halobacteria|o__Natrialbales|f__Natrialbaceae|g__Natrialba	0	0	0	0.19	0.23	0.27	0.15

k__Archaea|p__Euryarchaeota|c__Halobacteria|o__Natrialbales|f__Natrialbaceae|g__Natrialba|s__Natrialba_unclassified	0	0	0	0.19	0.23	0.27	0.15
k__Archaea|p__Euryarchaeota|c__Halobacteria|o__Natrialbales|f__Natrialbaceae|g__Natrinema	0	0	0	0.09	0.11	0.13	0.09
k__Archaea|p__Euryarchaeota|c__Halobacteria|o__Natrialbales|f__Natrialbaceae|g__Natrinema|s__Natrinema_altunense	0	0	0	0.06	0.01	0.01	0
k__Archaea|p__Euryarchaeota|c__Halobacteria|o__Natrialbales|f__Natrialbaceae|g__Natrinema|s__Natrinema_altunense|t__GCF_000337155	0	0	0	0.06	0.01	0.01	0
k__Archaea|p__Euryarchaeota|c__Halobacteria|o__Natrialbales|f__Natrialbaceae|g__Natrinema|s__Natrinema_gari	0	0	0	0.04	0.08	0.04	0.05
k__Archaea|p__Euryarchaeota|c__Halobacteria|o__Natrialbales|f__Natrialbaceae|g__Natrinema|s__Natrinema_gari|t__GCF_000337175	0	0	0	0.04	0.08	0.04	0.05
k__Archaea|p__Euryarchaeota|c__Halobacteria|o__Natrialbales|f__Natrialbaceae|g__Natrinema|s__Natrinema_unclassified	0	0	0	0	0.02	0.09	0.04
k__Archaea|p__Euryarchaeota|c__Halobacteria|o__Natrialbales|f__Natrialbaceae|g__Natronorubrum	0	0	0	0.06	0.08	0.04	0.02
k__Archaea|p__Euryarchaeota|c__Halobacteria|o__Natrialbales|f__Natrialbaceae|g__Natronorubrum|s__Natronorubrum_unclassified	0	0	0	0.06	0.08	0.04	0.02
k__Bacteria	87.32	96.66	94.06	5.94	4.35	2.6	2.01
k__Bacteria|p__Actinobacteria	15.19	13.32	3.08	0.08	0	0	0.24
k__Bacteria|p__Actinobacteria|c__Actinobacteria	15.19	13.32	3.08	0.08	0	0	0.24
k__Bacteria|p__Actinobacteria|c__Actinobacteria|o__Actinomycetales	11.42	11.86	1.44	0.08	0	0	0.15
k__Bacteria|p__Actinobacteria|c__Actinobacteria|o__Actinomycetales|f__Actinomycetaceae	0	0	0.03	0	0	0	0
k__Bacteria|p__Actinobacteria|c__Actinobacteria|o__Actinomycetales|f__Actinomycetaceae|g__Actinomyces	0	0	0.03	0	0	0	0
k__Bacteria|p__Actinobacteria|c__Actinobacteria|o__Actinomycetales|f__Actinomycetaceae|g__Actinomyces|s__Actinomyces_oris	0	0	0.03	0	0	0	0
k__Bacteria|p__Actinobacteria|c__Actinobacteria|o__Actinomycetales|f__Actinomycetaceae|g__Actinomyces|s__Actinomyces_oris|t_Actinomyces_oris_unclassified	0	0	0.03	0	0	0	0

k__Bacteria|p__Actinobacteria|c__Actinobacteria|o__Actinomycetales|f__Propionibacteriaceae	11.42	11.86	0.93	0.08	0	0	0.15
k__Bacteria|p__Actinobacteria|c__Actinobacteria|o__Actinomycetales|f__Propionibacteriaceae|g__Propionibacteriaceae_unclassified	11.42	11.86	0.93	0.08	0	0	0.15
k__Bacteria|p__Actinobacteria|c__Actinobacteria|o__Actinomycetales|f__Brevibacteriaceae	0	0	0.49	0	0	0	0
k__Bacteria|p__Actinobacteria|c__Actinobacteria|o__Actinomycetales|f__Brevibacteriaceae|g__Brevibacterium	0	0	0.49	0	0	0	0
k__Bacteria|p__Actinobacteria|c__Actinobacteria|o__Actinomycetales|f__Brevibacteriaceae|g__Brevibacterium|s__Brevibacterium_unclassified	0	0	0.49	0	0	0	0
k__Bacteria|p__Actinobacteria|c__Actinobacteria|o__Coriobacteriales	0	0.03	0	0	0	0	0
k__Bacteria|p__Actinobacteria|c__Actinobacteria|o__Coriobacteriales|f__Coriobacteriaceae	0	0.03	0	0	0	0	0
k__Bacteria|p__Actinobacteria|c__Actinobacteria|o__Coriobacteriales|f__Coriobacteriaceae|g__Atopobium	0	0.03	0	0	0	0	0
k__Bacteria|p__Actinobacteria|c__Actinobacteria|o__Coriobacteriales|f__Coriobacteriaceae|g__Atopobium|s__Atopobium_vaginae	0	0.03	0	0	0	0	0
k__Bacteria|p__Actinobacteria|c__Actinobacteria|o__Coriobacteriales|f__Coriobacteriaceae|g__Atopobium|s__Atopobium_vaginae|t__Atopobium_vaginae_unclassified	0	0.03	0	0	0	0	0
k__Bacteria|p__Actinobacteria|c__Actinobacteria|o__Corynebacteriales	3.63	1.35	1.35	0	0	0	0
k__Bacteria|p__Actinobacteria|c__Actinobacteria|o__Corynebacteriales|f__Corynebacteriaceae	3.63	1.35	1.35	0	0	0	0
k__Bacteria|p__Actinobacteria|c__Actinobacteria|o__Corynebacteriales|f__Corynebacteriaceae|g__Corynebacterium	3.63	1.35	1.35	0	0	0	0
k__Bacteria|p__Actinobacteria|c__Actinobacteria|o__Corynebacteriales|f__Corynebacteriaceae|g__Corynebacterium|s__Corynebacterium_jeikeium	0.3	0.24	0.35	0	0	0	0
k__Bacteria|p__Actinobacteria|c__Actinobacteria|o__Corynebacteriales|f__Corynebacteriaceae|g__Corynebacterium|s__Corynebacterium_jeikeium|t_Corynebacterium_jeikeium_unclassified	0.3	0.24	0.35	0	0	0	0
k__Bacteria|p__Actinobacteria|c__Actinobacteria|o__Corynebacteriales|f__Corynebacteriaceae|g__Corynebacterium|s__Corynebacterium_otitidis	3.33	1.11	1	0	0	0	0
k__Bacteria|p__Actinobacteria|c__Actinobacteria|o__Corynebacteriales|f__Corynebacteriaceae|g__Corynebacterium|s__Corynebacterium_otitidis|t_Corynebacterium_otitidis_unclassified	3.33	1.11	1	0	0	0	0

k__Bacteria|p__Actinobacteria|c__Actinobacteria|o__Micrococcales	0.14	0	0.08	0	0	0	0
k__Bacteria|p__Actinobacteria|c__Actinobacteria|o__Micrococcales|f__Micrococcaceae	0.14	0	0.08	0	0	0	0
k__Bacteria|p__Actinobacteria|c__Actinobacteria|o__Micrococcales|f__Micrococcaceae|g__Kocuria	0.14	0	0	0	0	0	0
k__Bacteria|p__Actinobacteria|c__Actinobacteria|o__Micrococcales|f__Micrococcaceae|g__Kocuria|s__Kocuria_palustris	0.14	0	0	0	0	0	0
k__Bacteria|p__Actinobacteria|c__Actinobacteria|o__Micrococcales|f__Micrococcaceae|g__Kocuria|s__Kocuria_palustris|t_Kocuria_palustris_unclassified	0.14	0	0	0	0	0	0
k__Bacteria|p__Actinobacteria|c__Actinobacteria|o__Micrococcales|f__Micrococcaceae|g__Rothia	0	0	0.08	0	0	0	0
k__Bacteria|p__Actinobacteria|c__Actinobacteria|o__Micrococcales|f__Micrococcaceae|g__Rothia|s__Rothia_mucilaginosa	0	0	0.08	0	0	0	0
k__Bacteria|p__Actinobacteria|c__Actinobacteria|o__Micrococcales|f__Micrococcaceae|g__Rothia|s__Rothia_mucilaginosa|t_Rothia_mucilaginosa_unclassified	0	0	0.08	0	0	0	0
k__Bacteria|p__Actinobacteria|c__Actinobacteria|o__Rubrobacterales	0	0.01	0.09	0	0	0	0
k__Bacteria|p__Actinobacteria|c__Actinobacteria|o__Rubrobacterales|f__Rubrobacteraceae	0	0.01	0.09	0	0	0	0
k__Bacteria|p__Actinobacteria|c__Actinobacteria|o__Rubrobacterales|f__Rubrobacteraceae|g__Rubrobacter	0	0.01	0.09	0	0	0	0
k__Bacteria|p__Actinobacteria|c__Actinobacteria|o__Rubrobacterales|f__Rubrobacteraceae|g__Rubrobacter|s__Rubrobacter_xylanophilus	0	0.01	0.09	0	0	0	0
k__Bacteria|p__Actinobacteria|c__Actinobacteria|o__Rubrobacterales|f__Rubrobacteraceae|g__Rubrobacter|s__Rubrobacter_xylanophilus|t__GCF_000014185	0	0.01	0.09	0	0	0	0
k__Bacteria|p__Actinobacteria|c__Actinobacteria|o__Solirubrobacterales	0	0.07	0	0	0	0	0
k__Bacteria|p__Actinobacteria|c__Actinobacteria|o__Solirubrobacterales|f__Solirubrobacterales_unclassified	0	0.07	0	0	0	0	0
k__Bacteria|p__Actinobacteria|c__Thermoleophilia	0	0	0.13	0	0	0	0
k__Bacteria|p__Actinobacteria|c__Thermoleophilia|o__Thermoleophilales	0	0	0.13	0	0	0	0

k__Bacteria|p__Actinobacteria|c__Thermoleophilia|o__Thermoleophilales|f__Thermoleophilaceae	0	0	0.13	0	0	0	0
k__Bacteria|p__Actinobacteria|c__Thermoleophilia|o__Thermoleophilales|f__Thermoleophilaceae|g__Thermoleophilum	0	0	0.13	0	0	0	0
k__Bacteria|p__Actinobacteria|c__Thermoleophilia|o__Thermoleophilales|f__Thermoleophilaceae|g__Thermoleophilum| s__Thermoleophilum_album	0	0	0.13	0	0	0	0
k__Bacteria|p__Actinobacteria|c__Thermoleophilia|o__Thermoleophilales|f__Thermoleophilaceae|g__Thermoleophilum| s__Thermoleophilum_album|t_Thermoleophilum_album_unclassified	0	0	0.13	0	0	0	0
k__Bacteria|p__Bacteroidetes	1.2	2.1	0.55	5.28	3.82	2.48	1.48
k__Bacteria|p__Bacteroidetes|c__Bacteroidetes_noname	0	0	0	5.28	3.82	2.48	1.04
k__Bacteria|p__Bacteroidetes|c__Bacteroidetes_noname|o__Bacteroidetes_Order_II_Incertae_sedis	0	0	0	5.28	3.82	2.48	1.04
k__Bacteria|p__Bacteroidetes|c__Bacteroidetes_noname|o__Bacteroidetes_Order_II_Incertae_sedis|f__Rhodothermaceae	0	0	0	5.28	3.82	2.48	1.04
k__Bacteria|p__Bacteroidetes|c__Bacteroidetes_noname|o__Bacteroidetes_Order_II_Incertae_sedis|f__Rhodothermaceae| g__Salinibacter	0	0	0	5.28	3.82	2.48	1.04
k__Bacteria|p__Bacteroidetes|c__Bacteroidetes_noname|o__Bacteroidetes_Order_II_Incertae_sedis|f__Rhodothermaceae| g__Salinibacter|s__Salinibacter_ruber	0	0	0	5.28	3.82	2.48	1.04
k__Bacteria|p__Bacteroidetes|c__Bacteroidetes_noname|o__Bacteroidetes_Order_II_Incertae_sedis|f__Rhodothermaceae| g__Salinibacter|s__Salinibacter_ruber|t__Salinibacter_ruber_unclassified	0	0	0	5.28	3.82	2.48	1.04
k__Bacteria|p__Bacteroidetes|c__Flavobacteriia	0.92	2.02	0.55	0	0	0	0.44
k__Bacteria|p__Bacteroidetes|c__Flavobacteriia|o__Flavobacteriales	0.92	2.02	0.55	0	0	0	0.44
k__Bacteria|p__Bacteroidetes|c__Flavobacteriia|o__Flavobacteriales|f__Flavobacteriaceae	0.92	2.02	0.55	0	0	0	0.44
k__Bacteria|p__Bacteroidetes|c__Flavobacteriia|o__Flavobacteriales|f__Flavobacteriaceae|g__Cloacibacterium	0.92	0	0.55	0	0	0	0
k__Bacteria|p__Bacteroidetes|c__Flavobacteriia|o__Flavobacteriales|f__Flavobacteriaceae|g__Cloacibacterium| s__Cloacibacterium_normanense	0.92	0	0.55	0	0	0	0
k__Bacteria|p__Bacteroidetes|c__Flavobacteriia|o__Flavobacteriales|f__Flavobacteriaceae|g__Cloacibacterium| s__Cloacibacterium_normanense|t_Cloacibacterium_normanense_unclassified	0.92	0	0.55	0	0	0	0
k__Bacteria|p__Bacteroidetes|c__Flavobacteriia|o__Flavobacteriales|f__Flavobacteriaceae|g__Chryseobacterium|	0	0	0	0	0	0	0

s__Chryseobacterium_unclassified							
k__Bacteria|p__Bacteroidetes|c__Flavobacteriia|o__Flavobacteriales|f__Flavobacteriaceae|g__Gillisia	0	0	0	0	0	0	0.44
k__Bacteria|p__Bacteroidetes|c__Flavobacteriia|o__Flavobacteriales|f__Flavobacteriaceae|g__Gillisia|s__Gillisia_unclassified	0	0	0	0	0	0	0.44
k__Bacteria|p__Bacteroidetes|c__Flavobacteriia|o__Flavobacteriales|f__Flavobacteriaceae|g__Riemerella	0	2.02	0	0	0	0	0
k__Bacteria|p__Bacteroidetes|c__Flavobacteriia|o__Flavobacteriales|f__Flavobacteriaceae|g__Riemerella| s__Riemerella_unclassified	0	2.02	0	0	0	0	0
k__Bacteria|p__Bacteroidetes|c__Sphingobacteriia	0.28	0.08	0	0	0	0	0
k__Bacteria|p__Bacteroidetes|c__Sphingobacteriia|o__Sphingobacteriales	0.28	0.08	0	0	0	0	0
k__Bacteria|p__Bacteroidetes|c__Sphingobacteriia|o__Sphingobacteriales|f__Sphingobacteriaceae	0.28	0.07	0	0	0	0	0
k__Bacteria|p__Bacteroidetes|c__Sphingobacteriia|o__Sphingobacteriales|f__Sphingobacteriaceae|g__Pedobacter	0.28	0.07	0	0	0	0	0
k__Bacteria|p__Bacteroidetes|c__Sphingobacteriia|o__Sphingobacteriales|f__Sphingobacteriaceae|g__Pedobacter| s__Pedobacter_unclassified	0.28	0.07	0	0	0	0	0
k__Bacteria|p__Bacteroidetes|c__Sphingobacteriia|o__Sphingobacteriales|f__Sphingobacteriales_noname	0	0	0	0	0	0	0
k__Bacteria|p__Bacteroidetes|c__Sphingobacteriia|o__Sphingobacteriales|f__Sphingobacteriales_noname|g__Niabella	0	0	0	0	0	0	0
k__Bacteria|p__Bacteroidetes|c__Sphingobacteriia|o__Sphingobacteriales|f__Sphingobacteriales_noname|g__Niabella| s__Niabella_unclassified	0	0	0	0	0	0	0
k__Bacteria|p__Deinococcus_Thermus	0.07	0.45	2.71	0	0	0	0.01
k__Bacteria|p__Deinococcus_Thermus|c__Deinococci	0.07	0.45	2.71	0	0	0	0.01
k__Bacteria|p__Deinococcus_Thermus|c__Deinococci|o__Deinococcales	0	0.07	0.03	0	0	0	0
k__Bacteria|p__Deinococcus_Thermus|c__Deinococci|o__Deinococcales|f__Deinococcaceae	0	0.07	0.03	0	0	0	0

k__Bacteria|p__Deinococcus_Thermus|c__Deinococci|o__Deinococcales|f__Deinococcaceae|g__Deinococcus	0	0.07	0.03	0	0	0	0
k__Bacteria|p__Deinococcus_Thermus|c__Deinococci|o__Deinococcales|f__Deinococcaceae|g__Deinococcus|s__Deinococcus_geothermalis	0	0	0.03	0	0	0	0
k__Bacteria|p__Deinococcus_Thermus|c__Deinococci|o__Deinococcales|f__Deinococcaceae|g__Deinococcus|s__Deinococcus_geothermalis|t__GCF_000196275	0	0	0.03	0	0	0	0
k__Bacteria|p__Deinococcus_Thermus|c__Deinococci|o__Deinococcales|f__Deinococcaceae|g__Deinococcus|s__Deinococcus_unclassified	0	0.07	0	0	0	0	0
k__Bacteria|p__Deinococcus_Thermus|c__Deinococci|o__Thermales	0.07	0.38	2.68	0	0	0	0.01
k__Bacteria|p__Deinococcus_Thermus|c__Deinococci|o__Thermales|f__Thermaceae	0.07	0.38	2.68	0	0	0	0.01
k__Bacteria|p__Deinococcus_Thermus|c__Deinococci|o__Thermales|f__Thermaceae|g__Meiothermus	0.07	0.23	2.68	0	0	0	0.01
k__Bacteria|p__Deinococcus_Thermus|c__Deinococci|o__Thermales|f__Thermaceae|g__Meiothermus|s__Meiothermus_silvanus	0.04	0.23	2.55	0	0	0	0.01
k__Bacteria|p__Deinococcus_Thermus|c__Deinococci|o__Thermales|f__Thermaceae|g__Meiothermus|s__Meiothermus_silvanus|t__GCF_000092125	0.04	0.23	2.55	0	0	0	0.01
k__Bacteria|p__Deinococcus_Thermus|c__Deinococci|o__Thermales|f__Thermaceae|g__Thermus	0.03	0.15	0.13	0	0	0	0
k__Bacteria|p__Deinococcus_Thermus|c__Deinococci|o__Thermales|f__Thermaceae|g__Thermus|s__Thermus_scotoductus	0.03	0.15	0.13	0	0	0	0
k__Bacteria|p__Deinococcus_Thermus|c__Deinococci|o__Thermales|f__Thermaceae|g__Thermus|s__Thermus_scotoductus|t__Thermus_scotoductus_unclassified	0.03	0.15	0.13	0	0	0	0
k__Bacteria|p__Firmicutes	6.09	0.85	5.24	0	0.15	0	0
k__Bacteria|p__Firmicutes|c__Bacilli	6.03	0.82	5.24	0	0.15	0	0
k__Bacteria|p__Firmicutes|c__Bacilli|o__Bacillales	5.34	0.21	3.25	0	0.15	0	0
k__Bacteria|p__Firmicutes|c__Bacilli|o__Bacillales|f__Bacillaceae	4.18	0	1.98	0	0.11	0	0
k__Bacteria|p__Firmicutes|c__Bacilli|o__Bacillales|f__Bacillaceae|g__Anoxybacillus	0.65	0	0	0	0	0	0

k__Bacteria|p__Firmicutes|c__Bacilli|o__Bacillales|f__Bacillaceae|g__Anoxybacillus|s__Anoxybacillus_flavithermus	0.59	0	0	0	0	0	0
k__Bacteria|p__Firmicutes|c__Bacilli|o__Bacillales|f__Bacillaceae|g__Anoxybacillus|s__Anoxybacillus_flavithermus|t__Anoxybacillus_flavithermus_unclassified	0.59	0	0	0	0	0	0
k__Bacteria|p__Firmicutes|c__Bacilli|o__Bacillales|f__Bacillaceae|g__Anoxybacillus|s__Anoxybacillus_unclassified	0.07	0	0	0	0	0	0
k__Bacteria|p__Firmicutes|c__Bacilli|o__Bacillales|f__Bacillaceae|g__Bacillus	3.53	0	1.98	0	0.11	0	0
k__Bacteria|p__Firmicutes|c__Bacilli|o__Bacillales|f__Bacillaceae|g__Bacillus|s__Bacillus_marmarensis	3.33	0	1.98	0	0.07	0	0
k__Bacteria|p__Firmicutes|c__Bacilli|o__Bacillales|f__Bacillaceae|g__Bacillus|s__Bacillus_marmarensis|t__GCF_000474275	3.33	0	1.98	0	0.07	0	0
k__Bacteria|p__Firmicutes|c__Bacilli|o__Bacillales|f__Bacillaceae|g__Bacillus|s__Bacillus_pseudofirmus	0	0	0	0	0.04	0	0
k__Bacteria|p__Firmicutes|c__Bacilli|o__Bacillales|f__Bacillaceae|g__Bacillus|s__Bacillus_pseudofirmus|t__GCF_000005825	0	0	0	0	0.04	0	0
k__Bacteria|p__Firmicutes|c__Bacilli|o__Bacillales|f__Bacillaceae|g__Bacillus|s__Bacillus_thermoamylovorans	0.19	0	0	0	0	0	0
k__Bacteria|p__Firmicutes|c__Bacilli|o__Bacillales|f__Bacillaceae|g__Bacillus|s__Bacillus_thermoamylovorans|t_Bacillus_thermoamylovorans__unclassified	0.19	0	0	0	0	0	0
k__Bacteria|p__Firmicutes|c__Bacilli|o__Bacillales|f__Bacillales_noname	0	0	0	0	0.05	0	0
k__Bacteria|p__Firmicutes|c__Bacilli|o__Bacillales|f__Bacillales_noname|g__Exiguobacterium	0	0	0	0	0.05	0	0
k__Bacteria|p__Firmicutes|c__Bacilli|o__Bacillales|f__Bacillales_noname|g__Exiguobacterium|s__Exiguobacterium_pavilionensis	0	0	0	0	0	0	0
k__Bacteria|p__Firmicutes|c__Bacilli|o__Bacillales|f__Bacillales_noname|g__Exiguobacterium|s__Exiguobacterium_pavilionensis|t__GCF_000416965	0	0	0	0	0	0	0
k__Bacteria|p__Firmicutes|c__Bacilli|o__Bacillales|f__Bacillales_noname|g__Exiguobacterium|s__Exiguobacterium_unclassified	0	0	0	0	0.04	0	0
k__Bacteria|p__Firmicutes|c__Bacilli|o__Bacillales|f__Staphylococcaceae	1.16	0.21	1.26	0	0	0	0
k__Bacteria|p__Firmicutes|c__Bacilli|o__Bacillales|f__Staphylococcaceae|g__Staphylococcus	1.16	0.21	1.26	0	0	0	0
k__Bacteria|p__Firmicutes|c__Bacilli|o__Bacillales|f__Staphylococcaceae|g__Staphylococcus|	0.38	0	0.03	0	0	0	0

s__Staphylococcus_auricularis							
k__Bacteria|p__Firmicutes|c__Bacilli|o__Bacillales|f__Staphylococcaceae|g__Staphylococcus|s__Staphylococcus_auricularis|t_Staphylococcus_auricularis_unclassified	0.38	0	0.03	0	0	0	0
k__Bacteria|p__Firmicutes|c__Bacilli|o__Bacillales|f__Staphylococcaceae|g__Staphylococcus|s__Staphylococcus_caprae_capitis	0.13	0.05	0	0	0	0	0
k__Bacteria|p__Firmicutes|c__Bacilli|o__Bacillales|f__Staphylococcaceae|g__Staphylococcus|s__Staphylococcus_caprae_capitis|t__Staphylococcus_caprae_capitis_unclassified	0.13	0.05	0	0	0	0	0
k__Bacteria|p__Firmicutes|c__Bacilli|o__Bacillales|f__Staphylococcaceae|g__Staphylococcus|s__Staphylococcus_hominis	0	0.04	0	0	0	0	0
k__Bacteria|p__Firmicutes|c__Bacilli|o__Bacillales|f__Staphylococcaceae|g__Staphylococcus|s__Staphylococcus_hominis|t__Staphylococcus_hominis_unclassified	0	0.04	0	0	0	0	0
k__Bacteria|p__Firmicutes|c__Bacilli|o__Bacillales|f__Staphylococcaceae|g__Staphylococcus|s__Staphylococcus_warneri	0.65	0.11	1.23	0	0	0	0
k__Bacteria|p__Firmicutes|c__Bacilli|o__Bacillales|f__Staphylococcaceae|g__Staphylococcus|s__Staphylococcus_warneri|t__Staphylococcus_warneri_unclassified	0.65	0.11	1.23	0	0	0	0
k__Bacteria|p__Firmicutes|c__Bacilli|o__Lactobacillales	0.69	0.61	1.99	0	0	0	0
k__Bacteria|p__Firmicutes|c__Bacilli|o__Lactobacillales|f__Aerococcaceae	0	0	0.07	0	0	0	0
k__Bacteria|p__Firmicutes|c__Bacilli|o__Lactobacillales|f__Aerococcaceae|g__Aerococcus	0	0	0.07	0	0	0	0
k__Bacteria|p__Firmicutes|c__Bacilli|o__Lactobacillales|f__Aerococcaceae|g__Aerococcus|s__Aerococcus_viridans	0	0	0.07	0	0	0	0
k__Bacteria|p__Firmicutes|c__Bacilli|o__Lactobacillales|f__Aerococcaceae|g__Aerococcus|s__Aerococcus_viridans|t__Aerococcus_viridans_unclassified	0	0	0.07	0	0	0	0
k__Bacteria|p__Firmicutes|c__Bacilli|o__Lactobacillales|f__Carnobacteriaceae	0.02	0.08	0.01	0	0	0	0
k__Bacteria|p__Firmicutes|c__Bacilli|o__Lactobacillales|f__Carnobacteriaceae|g__Alloiococcus	0.02	0.08	0.01	0	0	0	0
k__Bacteria|p__Firmicutes|c__Bacilli|o__Lactobacillales|f__Carnobacteriaceae|g__Alloiococcus|s__Alloiococcus_otitis	0.02	0.08	0.01	0	0	0	0
k__Bacteria|p__Firmicutes|c__Bacilli|o__Lactobacillales|f__Carnobacteriaceae|g__Alloiococcus|s__Alloiococcus_otitis|t_Alloiococcus_otitis_unclassified	0.02	0.08	0.01	0	0	0	0

k__Bacteria|p__Firmicutes|c__Bacilli|o__Lactobacillales|f__Lactobacillaceae	0	0	1.23	0	0	0	0
k__Bacteria|p__Firmicutes|c__Bacilli|o__Lactobacillales|f__Lactobacillaceae|g__Lactobacillus	0	0	1.23	0	0	0	0
k__Bacteria|p__Firmicutes|c__Bacilli|o__Lactobacillales|f__Lactobacillaceae|g__Lactobacillus|s__Lactobacillus_jensenii	0	0	1.23	0	0	0	0
k__Bacteria|p__Firmicutes|c__Bacilli|o__Lactobacillales|f__Lactobacillaceae|g__Lactobacillus|s__Lactobacillus_jensenii|t_Lactobacillus_jensenii_unclassified	0	0	1.23	0	0	0	0
k__Bacteria|p__Firmicutes|c__Bacilli|o__Lactobacillales|f__Leuconostocaceae	0.41	0.43	0.3	0	0	0	0
k__Bacteria|p__Firmicutes|c__Bacilli|o__Lactobacillales|f__Leuconostocaceae|g__Leuconostoc	0.41	0.43	0.3	0	0	0	0
k__Bacteria|p__Firmicutes|c__Bacilli|o__Lactobacillales|f__Leuconostocaceae|g__Leuconostoc|s__Leuconostoc_gasicomitatum	0	0.05	0	0	0	0	0
k__Bacteria|p__Firmicutes|c__Bacilli|o__Lactobacillales|f__Leuconostocaceae|g__Leuconostoc|s__Leuconostoc_gasicomitatum|t__GCF_000196855	0	0.05	0	0	0	0	0
k__Bacteria|p__Firmicutes|c__Bacilli|o__Lactobacillales|f__Leuconostocaceae|g__Leuconostoc|s__Leuconostoc_gelidum	0	0.23	0.3	0	0	0	0
k__Bacteria|p__Firmicutes|c__Bacilli|o__Lactobacillales|f__Leuconostocaceae|g__Leuconostoc|s__Leuconostoc_gelidum|t_Leuconostoc_gelidum_unclassified	0	0.23	0.3	0	0	0	0
k__Bacteria|p__Firmicutes|c__Bacilli|o__Lactobacillales|f__Leuconostocaceae|g__Leuconostoc|s__Leuconostoc_unclassified	0.41	0.14	0	0	0	0	0
k__Bacteria|p__Firmicutes|c__Bacilli|o__Lactobacillales|f__Streptococcaceae	0.25	0.11	0.39	0	0	0	0
k__Bacteria|p__Firmicutes|c__Bacilli|o__Lactobacillales|f__Streptococcaceae|g__Lactococcus	0	0	0.37	0	0	0	0
k__Bacteria|p__Firmicutes|c__Bacilli|o__Lactobacillales|f__Streptococcaceae|g__Lactococcus|s__Lactococcus_lactis	0	0	0.37	0	0	0	0
k__Bacteria|p__Firmicutes|c__Bacilli|o__Lactobacillales|f__Streptococcaceae|g__Lactococcus|s__Lactococcus_lactis|t_Lactococcus_lactis_unclassified	0	0	0.37	0	0	0	0
k__Bacteria|p__Firmicutes|c__Bacilli|o__Lactobacillales|f__Streptococcaceae|g__Streptococcus	0.25	0.11	0.02	0	0	0	0
k__Bacteria|p__Firmicutes|c__Bacilli|o__Lactobacillales|f__Streptococcaceae|g__Streptococcus|s__Streptococcus_australis	0	0	0.02	0	0	0	0
k__Bacteria|p__Firmicutes|c__Bacilli|o__Lactobacillales|f__Streptococcaceae|g__Streptococcus|	0	0	0.02	0	0	0	0

s__Streptococcus_australis|t__Streptococcus_australis_unclassified							
k__Bacteria|p__Firmicutes|c__Bacilli|o__Lactobacillales|f__Streptococcaceae|g__Streptococcus|s__Streptococcus_mitis	0.25	0	0	0	0	0	0
k__Bacteria|p__Firmicutes|c__Bacilli|o__Lactobacillales|f__Streptococcaceae|g__Streptococcus|s__Streptococcus_mitis|t__Streptococcus_mitis__unclassified	0.25	0	0	0	0	0	0
k__Bacteria|p__Firmicutes|c__Bacilli|o__Lactobacillales|f__Streptococcaceae|g__Streptococcus|s__Streptococcus_suis	0	0.08	0	0	0	0	0
k__Bacteria|p__Firmicutes|c__Bacilli|o__Lactobacillales|f__Streptococcaceae|g__Streptococcus|s__Streptococcus_suis|t__Streptococcus_suis_unclassified	0	0.08	0	0	0	0	0
k__Bacteria|p__Firmicutes|c__Bacilli|o__Lactobacillales|f__Streptococcaceae|g__Streptococcus|s__Streptococcus_thermophilus	0	0.02	0	0	0	0	0
k__Bacteria|p__Firmicutes|c__Bacilli|o__Lactobacillales|f__Streptococcaceae|g__Streptococcus|s__Streptococcus_thermophilus|t__Streptococcus_thermophilus_unclassified	0	0.02	0	0	0	0	0
k__Bacteria|p__Firmicutes|c__Clostridia	0.07	0.03	0	0	0	0	0
k__Bacteria|p__Firmicutes|c__Clostridia|o__Clostridiales	0.04	0.03	0	0	0	0	0
k__Bacteria|p__Firmicutes|c__Clostridia|o__Clostridiales|f__Clostridiales_Family_XI_Incertae_Sedis	0.04	0.03	0	0	0	0	0
k__Bacteria|p__Firmicutes|c__Clostridia|o__Clostridiales|f__Clostridiales_Family_XI_Incertae_Sedis|g__Finegoldia	0.04	0.02	0	0	0	0	0
k__Bacteria|p__Firmicutes|c__Clostridia|o__Clostridiales|f__Clostridiales_Family_XI_Incertae_Sedis|g__Finegoldia|s__Finegoldia_magna	0.04	0.02	0	0	0	0	0
k__Bacteria|p__Firmicutes|c__Clostridia|o__Clostridiales|f__Clostridiales_Family_XI_Incertae_Sedis|g__Finegoldia|s__Finegoldia_magna|t__Finegoldia_magna_unclassified	0.04	0.02	0	0	0	0	0
k__Bacteria|p__Firmicutes|c__Clostridia|o__Clostridiales|f__Clostridiales_Family_XI_Incertae_Sedis|g__Peptoniphilus	0	0.01	0	0	0	0	0
k__Bacteria|p__Firmicutes|c__Clostridia|o__Clostridiales|f__Clostridiales_Family_XI_Incertae_Sedis|g__Peptoniphilus|s__Peptoniphilus_rhinitidis	0	0.01	0	0	0	0	0
k__Bacteria|p__Firmicutes|c__Clostridia|o__Clostridiales|f__Clostridiales_Family_XI_Incertae_Sedis|g__Peptoniphilus|s__Peptoniphilus_rhinitidis|t__GCF_000246925	0	0.01	0	0	0	0	0
k__Bacteria|p__Firmicutes|c__Clostridia|o__Thermoanaerobacterales	0.03	0	0	0	0	0	0
k__Bacteria|p__Firmicutes|c__Clostridia|o__Thermoanaerobacterales|	0.03	0	0	0	0	0	0

f__Thermoanaerobacterales_Family_III_Incertae_Sedis							
k__Bacteria|p__Firmicutes|c__Clostridia|o__Thermoanaerobacterales|f__Thermoanaerobacterales_Family_III_Incertae_Sedis| g__Thermoanaerobacterium	0.03	0	0	0	0	0	0
k__Bacteria|p__Firmicutes|c__Clostridia|o__Thermoanaerobacterales|f__Thermoanaerobacterales_Family_III_Incertae_Sedis| g__Thermoanaerobacterium|s__Thermoanaerobacterium_unclassified	0.03	0	0	0	0	0	0
k__Bacteria|p__Firmicutes|c__Negativicutes	0	0	0	0	0	0	0
k__Bacteria|p__Firmicutes|c__Negativicutes|o__Selenomonadales	0	0	0	0	0	0	0
k__Bacteria|p__Firmicutes|c__Negativicutes|o__Selenomonadales|f__Veillonellaceae	0	0	0	0	0	0	0
k__Bacteria|p__Firmicutes|c__Negativicutes|o__Selenomonadales|f__Veillonellaceae|g__Veillonella	0	0	0	0	0	0	0
k__Bacteria|p__Firmicutes|c__Negativicutes|o__Selenomonadales|f__Veillonellaceae|g__Veillonella| s__Veillonella_unclassified	0	0	0	0	0	0	0
k__Bacteria|p__Proteobacteria	62.8	79.55	87.93	0.59	0.38	0.25	0.34
k__Bacteria|p__Proteobacteria|c__Alphaproteobacteria	4.1	5.51	17.73	0.05	0.02	0	0.07
k__Bacteria|p__Proteobacteria|c__Alphaproteobacteria|o__Caulobacterales	1.08	0.2	1.08	0	0	0	0.02
k__Bacteria|p__Proteobacteria|c__Alphaproteobacteria|o__Caulobacterales|f__Caulobacteraceae	1.08	0.2	1.08	0	0	0	0.02
k__Bacteria|p__Proteobacteria|c__Alphaproteobacteria|o__Caulobacterales|f__Caulobacteraceae|g__Asticcacaulis	0.21	0	0	0	0	0	0
k__Bacteria|p__Proteobacteria|c__Alphaproteobacteria|o__Caulobacterales|f__Caulobacteraceae|g__Asticcacaulis| s__Asticcacaulis_unclassified	0.21	0	0	0	0	0	0
k__Bacteria|p__Proteobacteria|c__Alphaproteobacteria|o__Caulobacterales|f__Caulobacteraceae|g__Brevundimonas	0.87	0.08	0.42	0	0	0	0.02
k__Bacteria|p__Proteobacteria|c__Alphaproteobacteria|o__Caulobacterales|f__Caulobacteraceae|g__Brevundimonas| s__Brevundimonas_unclassified	0.87	0.08	0.42	0	0	0	0.02
k__Bacteria|p__Proteobacteria|c__Alphaproteobacteria|o__Caulobacterales|f__Caulobacteraceae|g__Caulobacter	0	0.13	0.66	0	0	0	0

k__Bacteria|p__Proteobacteria|c__Alphaproteobacteria|o__Caulobacterales|f__Caulobacteraceae|g__Caulobacter|s__Caulobacter_unclassified	0	0.13	0.66	0	0	0	0
k__Bacteria|p__Proteobacteria|c__Alphaproteobacteria|o__Rhizobiales	2.4	4.81	13.7	0	0	0	0.03
k__Bacteria|p__Proteobacteria|c__Alphaproteobacteria|o__Rhizobiales|f__Bradyrhizobiaceae	1.93	4.49	10.88	0	0	0	0.03
k__Bacteria|p__Proteobacteria|c__Alphaproteobacteria|o__Rhizobiales|f__Bradyrhizobiaceae|g__Afipia	0	0.01	0	0	0	0	0
k__Bacteria|p__Proteobacteria|c__Alphaproteobacteria|o__Rhizobiales|f__Bradyrhizobiaceae|g__Afipia|s__Afipia_unclassified	0	0.01	0	0	0	0	0
k__Bacteria|p__Proteobacteria|c__Alphaproteobacteria|o__Rhizobiales|f__Bradyrhizobiaceae|g__Bradyrhizobium	1.93	4.04	9.42	0	0	0	0
k__Bacteria|p__Proteobacteria|c__Alphaproteobacteria|o__Rhizobiales|f__Bradyrhizobiaceae|g__Bradyrhizobium|s__Bradyrhizobium_japonicum	0	0.02	0.01	0	0	0	0
k__Bacteria|p__Proteobacteria|c__Alphaproteobacteria|o__Rhizobiales|f__Bradyrhizobiaceae|g__Bradyrhizobium|s__Bradyrhizobium_japonicum|t__Bradyrhizobium_japonicum_unclassified	0	0.02	0.01	0	0	0	0
k__Bacteria|p__Proteobacteria|c__Alphaproteobacteria|o__Rhizobiales|f__Bradyrhizobiaceae|g__Bradyrhizobium|s__Bradyrhizobium_sp_DFCI_1	1.93	4.02	9.42	0	0	0	0
k__Bacteria|p__Proteobacteria|c__Alphaproteobacteria|o__Rhizobiales|f__Bradyrhizobiaceae|g__Bradyrhizobium|s__Bradyrhizobium_sp_DFCI_1|t__GCF_000465325	1.93	4.02	9.42	0	0	0	0
k__Bacteria|p__Proteobacteria|c__Alphaproteobacteria|o__Rhizobiales|f__Bradyrhizobiaceae|g__Rhodopseudomonas	0	0.45	1.46	0	0	0	0
k__Bacteria|p__Proteobacteria|c__Alphaproteobacteria|o__Rhizobiales|f__Bradyrhizobiaceae|g__Rhodopseudomonas|s__Rhodopseudomonas_palustris	0	0.45	1.46	0	0	0	0
k__Bacteria|p__Proteobacteria|c__Alphaproteobacteria|o__Rhizobiales|f__Bradyrhizobiaceae|g__Rhodopseudomonas|s__Rhodopseudomonas_palustris|t__Rhodopseudomonas_palustris_unclassified	0	0.45	1.46	0	0	0	0
k__Bacteria|p__Proteobacteria|c__Alphaproteobacteria|o__Rhizobiales|f__Hyphomicrobiaceae	0.31	0.28	1.27	0	0	0	0.03
k__Bacteria|p__Proteobacteria|c__Alphaproteobacteria|o__Rhizobiales|f__Hyphomicrobiaceae|g__Hyphomicrobiaceae_unclassified	0.31	0.27	1.27	0	0	0	0.03
k__Bacteria|p__Proteobacteria|c__Alphaproteobacteria|o__Rhizobiales|f__Hyphomicrobiaceae|g__Hyphomicrobium	0	0	0	0	0	0	0
k__Bacteria|p__Proteobacteria|c__Alphaproteobacteria|o__Rhizobiales|f__Hyphomicrobiaceae|g__Hyphomicrobium|s__Hyphomicrobium_unclassified	0	0	0	0	0	0	0

k__Bacteria|p__Proteobacteria|c__Alphaproteobacteria|o__Rhizobiales|f__Methylobacteriaceae	0	0.01	0.69	0	0	0	0
k__Bacteria|p__Proteobacteria|c__Alphaproteobacteria|o__Rhizobiales|f__Methylobacteriaceae|g__Methylobacterium	0	0.01	0.69	0	0	0	0
k__Bacteria|p__Proteobacteria|c__Alphaproteobacteria|o__Rhizobiales|f__Methylobacteriaceae|g__Methylobacterium|s__Methylobacterium_organophilum	0	0	0.26	0	0	0	0
k__Bacteria|p__Proteobacteria|c__Alphaproteobacteria|o__Rhizobiales|f__Methylobacteriaceae|g__Methylobacterium|s__Methylobacterium_organophilum|t_Methylobacterium_organophilum_unclassified	0	0	0.26	0	0	0	0
k__Bacteria|p__Proteobacteria|c__Alphaproteobacteria|o__Rhizobiales|f__Methylobacteriaceae|g__Methylobacterium|s__Methylobacterium_radiotolerans	0	0.01	0.44	0	0	0	0
k__Bacteria|p__Proteobacteria|c__Alphaproteobacteria|o__Rhizobiales|f__Methylobacteriaceae|g__Methylobacterium|s__Methylobacterium_radiotolerans|t_Methylobacterium_radiotolerans_unclassified	0	0.01	0.44	0	0	0	0
k__Bacteria|p__Proteobacteria|c__Alphaproteobacteria|o__Rhizobiales|f__Rhizobiaceae	0.16	0	0	0	0	0	0
k__Bacteria|p__Proteobacteria|c__Alphaproteobacteria|o__Rhizobiales|f__Rhizobiaceae|g__Agrobacterium	0.16	0	0	0	0	0	0
k__Bacteria|p__Proteobacteria|c__Alphaproteobacteria|o__Rhizobiales|f__Rhizobiaceae|g__Agrobacterium|s__Agrobacterium_salinitolerans	0.16	0	0	0	0	0	0
k__Bacteria|p__Proteobacteria|c__Alphaproteobacteria|o__Rhizobiales|f__Rhizobiaceae|g__Agrobacterium|s__Agrobacterium_salinitolerans|t_Agrobacterium_salinitolerans_unclassified	0.16	0	0	0	0	0	0
k__Bacteria|p__Proteobacteria|c__Alphaproteobacteria|o__Rhizobiales|f__Rhodobiaceae	0	0.04	0.86	0	0	0	0
k__Bacteria|p__Proteobacteria|c__Alphaproteobacteria|o__Rhizobiales|f__Rhodobiaceae|g__Rhodobiaceae_unclassified	0	0.04	0.86	0	0	0	0
k__Bacteria|p__Proteobacteria|c__Alphaproteobacteria|o__Rhodobacterales	0.17	0.01	0	0	0.02	0	0
k__Bacteria|p__Proteobacteria|c__Alphaproteobacteria|o__Rhodobacterales|f__Rhodobacteraceae	0.17	0.01	0	0	0.02	0	0
k__Bacteria|p__Proteobacteria|c__Alphaproteobacteria|o__Rhodobacterales|f__Rhodobacteraceae|g__Paracoccus	0.17	0.01	0	0	0.02	0	0
k__Bacteria|p__Proteobacteria|c__Alphaproteobacteria|o__Rhodobacterales|f__Rhodobacteraceae|g__Paracoccus|s__Paracoccus_denitrificans	0	0	0	0	0	0	0
k__Bacteria|p__Proteobacteria|c__Alphaproteobacteria|o__Rhodobacterales|f__Rhodobacteraceae|g__Paracoccus|s__Paracoccus_denitrificans|t__Paracoccus_denitrificans_unclassified	0	0	0	0	0	0	0
k__Bacteria|p__Proteobacteria|c__Alphaproteobacteria|o__Rhodobacterales|f__Rhodobacteraceae|g__Paracoccus|	0.17	0	0	0	0.02	0	0

s__Paracoccus_unclassified							
k__Bacteria|p__Proteobacteria|c__Alphaproteobacteria|o__Sphingomonadales	0.45	0.49	2.95	0.05	0	0	0.01
k__Bacteria|p__Proteobacteria|c__Alphaproteobacteria|o__Sphingomonadales|f__Sphingomonadaceae	0.45	0.49	2.95	0.05	0	0	0.01
k__Bacteria|p__Proteobacteria|c__Alphaproteobacteria|o__Sphingomonadales|f__Sphingomonadaceae|g__Sphingobium	0	0	0	0.05	0	0	0
k__Bacteria|p__Proteobacteria|c__Alphaproteobacteria|o__Sphingomonadales|f__Sphingomonadaceae|g__Sphingobium|s__Sphingobium_unclassified	0	0	0	0.05	0	0	0
k__Bacteria|p__Proteobacteria|c__Alphaproteobacteria|o__Sphingomonadales|f__Sphingomonadaceae|g__Sphingomonas	0.45	0.49	2.95	0	0	0	0.01
k__Bacteria|p__Proteobacteria|c__Alphaproteobacteria|o__Sphingomonadales|f__Sphingomonadaceae|g__Sphingomonas|s__Sphingomonas_echinoides	0.45	0.49	2.95	0	0	0	0.01
k__Bacteria|p__Proteobacteria|c__Alphaproteobacteria|o__Sphingomonadales|f__Sphingomonadaceae|g__Sphingomonas|s__Sphingomonas_echinoides|t__GCF_000241465	0.45	0.49	2.95	0	0	0	0.01
k__Bacteria|p__Proteobacteria|c__Betaproteobacteria	37.77	30.97	48.14	0.13	0	0	0.01
k__Bacteria|p__Proteobacteria|c__Betaproteobacteria|o__Burkholderiales	37.77	30.97	47.98	0.13	0	0	0.01
k__Bacteria|p__Proteobacteria|c__Betaproteobacteria|o__Burkholderiales|f__Burkholderiaceae	37.29	15.15	46.68	0.13	0	0	0
k__Bacteria|p__Proteobacteria|c__Betaproteobacteria|o__Burkholderiales|f__Burkholderiaceae|g__Burkholderia	0	12.51	0	0.13	0	0	0
k__Bacteria|p__Proteobacteria|c__Betaproteobacteria|o__Burkholderiales|f__Burkholderiaceae|g__Burkholderia|s__Burkholderia_unclassified	0	12.51	0	0.13	0	0	0
k__Bacteria|p__Proteobacteria|c__Betaproteobacteria|o__Burkholderiales|f__Burkholderiaceae|g__Cupriavidus	0	0.07	0	0	0	0	0
k__Bacteria|p__Proteobacteria|c__Betaproteobacteria|o__Burkholderiales|f__Burkholderiaceae|g__Cupriavidus|s__Cupriavidus_unclassified	0	0.07	0	0	0	0	0
k__Bacteria|p__Proteobacteria|c__Betaproteobacteria|o__Burkholderiales|f__Burkholderiaceae|g__Limnobacter	0.27	0	1.63	0	0	0	0
k__Bacteria|p__Proteobacteria|c__Betaproteobacteria|o__Burkholderiales|f__Burkholderiaceae|g__Limnobacter|s__Limnobacter_sp_CACIAM_66H1	0.27	0	1.63	0	0	0	0

k__Bacteria|p__Proteobacteria|c__Betaproteobacteria|o__Burkholderiales|f__Burkholderiaceae|g__Paraburkholderia	32.98	0	37.7	0	0	0	0
k__Bacteria|p__Proteobacteria|c__Betaproteobacteria|o__Burkholderiales|f__Burkholderiaceae|g__Paraburkholderia|s__Paraburkholderia_fungorum	32.98	0	37.7	0	0	0	0
k__Bacteria|p__Proteobacteria|c__Betaproteobacteria|o__Burkholderiales|f__Burkholderiaceae|g__Paraburkholderia|s__Paraburkholderia_fungorum|t__Paraburkholderia_fungorum_unclassified	32.98	0	37.7	0	0	0	0
k__Bacteria|p__Proteobacteria|c__Betaproteobacteria|o__Burkholderiales|f__Burkholderiaceae|g__Ralstonia	4.02	2.57	7.35	0	0	0	0
k__Bacteria|p__Proteobacteria|c__Betaproteobacteria|o__Burkholderiales|f__Burkholderiaceae|g__Ralstonia|s__Ralstonia_mannitolilytica	0	0	0.07	0	0	0	0
k__Bacteria|p__Proteobacteria|c__Betaproteobacteria|o__Burkholderiales|f__Burkholderiaceae|g__Ralstonia|s__Ralstonia_mannitolilytica|t_Ralstonia_mannitolilytica_unclassified	0	0	0.07	0	0	0	0
k__Bacteria|p__Proteobacteria|c__Betaproteobacteria|o__Burkholderiales|f__Burkholderiaceae|g__Ralstonia|s__Ralstonia_pickettii	2.25	0.15	3.92	0	0	0	0
k__Bacteria|p__Proteobacteria|c__Betaproteobacteria|o__Burkholderiales|f__Burkholderiaceae|g__Ralstonia|s__Ralstonia_pickettii|t__Ralstonia_pickettii_unclassified	2.25	0.15	3.92	0	0	0	0
k__Bacteria|p__Proteobacteria|c__Betaproteobacteria|o__Burkholderiales|f__Burkholderiaceae|g__Ralstonia|s__Ralstonia_unclassified	1.79	2.42	3.36	0	0	0	0
k__Bacteria|p__Proteobacteria|c__Betaproteobacteria|o__Burkholderiales|f__Comamonadaceae	0.49	15.79	1.3	0	0	0	0.01
k__Bacteria|p__Proteobacteria|c__Betaproteobacteria|o__Burkholderiales|f__Comamonadaceae|g__Acidovorax	0.05	6.21	0	0	0	0	0
k__Bacteria|p__Proteobacteria|c__Betaproteobacteria|o__Burkholderiales|f__Comamonadaceae|g__Acidovorax|s__Acidovorax_ebreus	0.05	6.21	0	0	0	0	0
k__Bacteria|p__Proteobacteria|c__Betaproteobacteria|o__Burkholderiales|f__Comamonadaceae|g__Acidovorax|s__Acidovorax_ebreus|t__GCF_000022305	0.05	6.21	0	0	0	0	0
k__Bacteria|p__Proteobacteria|c__Betaproteobacteria|o__Burkholderiales|f__Comamonadaceae|g__Alicycliphilus	0	9.29	0.16	0	0	0	0
k__Bacteria|p__Proteobacteria|c__Betaproteobacteria|o__Burkholderiales|f__Comamonadaceae|g__Alicycliphilus|s__Alicycliphilus_denitrificans	0	0.01	0	0	0	0	0
k__Bacteria|p__Proteobacteria|c__Betaproteobacteria|o__Burkholderiales|f__Comamonadaceae|g__Alicycliphilus|s__Alicycliphilus_denitrificans|t__Alicycliphilus_denitrificans_unclassified	0	0.01	0	0	0	0	0
k__Bacteria|p__Proteobacteria|c__Betaproteobacteria|o__Burkholderiales|f__Comamonadaceae|g__Alicycliphilus|s__Alicycliphilus_unclassified	0	9.28	0.16	0	0	0	0

k__Bacteria|p__Proteobacteria|c__Betaproteobacteria|o__Burkholderiales|f__Comamonadaceae|g__Comamonas	0	0.14	0	0	0	0	0
k__Bacteria|p__Proteobacteria|c__Betaproteobacteria|o__Burkholderiales|f__Comamonadaceae|g__Comamonas|s__Comamonas_unclassified	0	0.14	0	0	0	0	0
k__Bacteria|p__Proteobacteria|c__Betaproteobacteria|o__Burkholderiales|f__Comamonadaceae|g__Delftia	0.36	0.01	0.86	0	0	0	0.01
k__Bacteria|p__Proteobacteria|c__Betaproteobacteria|o__Burkholderiales|f__Comamonadaceae|g__Delftia|s__Delftia_acidovorans	0.07	0	0.06	0	0	0	0
k__Bacteria|p__Proteobacteria|c__Betaproteobacteria|o__Burkholderiales|f__Comamonadaceae|g__Delftia|s__Delftia_acidovorans|t__Delftia_acidovorans_unclassified	0.07	0	0.06	0	0	0	0
k__Bacteria|p__Proteobacteria|c__Betaproteobacteria|o__Burkholderiales|f__Comamonadaceae|g__Delftia|s__Delftia_tsuruhatensis	0.18	0	0.8	0	0	0	0
k__Bacteria|p__Proteobacteria|c__Betaproteobacteria|o__Burkholderiales|f__Comamonadaceae|g__Delftia|s__Delftia_tsuruhatensis|t_Delftia_tsuruhatensis_unclassified	0.18	0	0.8	0	0	0	0
k__Bacteria|p__Proteobacteria|c__Betaproteobacteria|o__Burkholderiales|f__Comamonadaceae|g__Delftia|s__Delftia_unclassified	0.11	0.01	0.23	0	0	0	0.01
k__Bacteria|p__Proteobacteria|c__Betaproteobacteria|o__Burkholderiales|f__Comamonadaceae|g__Limnohabitans	0	0	0	0	0	0	0
k__Bacteria|p__Proteobacteria|c__Betaproteobacteria|o__Burkholderiales|f__Comamonadaceae|g__Limnohabitans|s__Limnohabitans_unclassified	0	0	0	0	0	0	0
k__Bacteria|p__Proteobacteria|c__Betaproteobacteria|o__Burkholderiales|f__Comamonadaceae|g__Polaromonas	0	0.09	0	0	0	0	0
k__Bacteria|p__Proteobacteria|c__Betaproteobacteria|o__Burkholderiales|f__Comamonadaceae|g__Polaromonas|s__Polaromonas_unclassified	0	0.09	0	0	0	0	0
k__Bacteria|p__Proteobacteria|c__Betaproteobacteria|o__Burkholderiales|f__Comamonadaceae||g__Leptothrix	0.03	0	0.03	0	0	0	0
k__Bacteria|p__Proteobacteria|c__Betaproteobacteria|o__Burkholderiales|f__Comamonadaceae||g__Leptothrix|s__Leptothrix_ochracea	0.03	0	0.03	0	0	0	0
k__Bacteria|p__Proteobacteria|c__Betaproteobacteria|o__Burkholderiales|f__Comamonadaceae||g__Leptothrix|s__Leptothrix_ochracea|t__Leptothrix_ochracea_unclassified	0.03	0	0.03	0	0	0	0
k__Bacteria|p__Proteobacteria|c__Betaproteobacteria|o__Burkholderiales|f__Comamonadaceae|g__Variovorax	0.05	0	0.02	0	0	0	0.01
k__Bacteria|p__Proteobacteria|c__Betaproteobacteria|o__Burkholderiales|f__Comamonadaceae|g__Variovorax|s__Variovorax_unclassified	0.05	0	0.02	0	0	0	0.01

k__Bacteria|p__Proteobacteria|c__Betaproteobacteria|o__Burkholderiales|f__Oxalobacteraceae	0	0.03	0	0	0	0	0
k__Bacteria|p__Proteobacteria|c__Betaproteobacteria|o__Burkholderiales|f__Oxalobacteraceae|g__Janthinobacterium	0	0.03	0	0	0	0	0
k__Bacteria|p__Proteobacteria|c__Betaproteobacteria|o__Burkholderiales|f__Oxalobacteraceae|g__Janthinobacterium| s__Janthinobacterium_unclassified	0	0.03	0	0	0	0	0
k__Bacteria|p__Proteobacteria|c__Betaproteobacteria|o__Rhodocyclales	0	0	0.16	0	0	0	0
k__Bacteria|p__Proteobacteria|c__Betaproteobacteria|o__Rhodocyclales|f__Rhodocyclaceae	0	0	0.16	0	0	0	0
k__Bacteria|p__Proteobacteria|c__Betaproteobacteria|o__Rhodocyclales|f__Rhodocyclaceae|g__Methyloversatilis	0	0	0.16	0	0	0	0
k__Bacteria|p__Proteobacteria|c__Betaproteobacteria|o__Rhodocyclales|f__Rhodocyclaceae|g__Methyloversatilis| s__Methyloversatilis_unclassified	0	0	0.16	0	0	0	0
k__Bacteria|p__Proteobacteria|c__Betaproteobacteria|o__Rhodocyclales|f__Rhodocyclaceae|g__Methyloversatilis| s__Methyloversatilis_universalis	0	0	0	0	0	0	0
k__Bacteria|p__Proteobacteria|c__Betaproteobacteria|o__Rhodocyclales|f__Rhodocyclaceae|g__Methyloversatilis| s__Methyloversatilis_universalis|t__Methyloversatilis_universalis_unclassified	0	0	0	0	0	0	0
k__Bacteria|p__Proteobacteria|c__Gammaproteobacteria	21.88	43.08	14.03	0.41	0.36	0.25	0.26
k__Bacteria|p__Proteobacteria|c__Gammaproteobacteria|o__Aeromonadales	0.93	0	0	0	0	0	0
k__Bacteria|p__Proteobacteria|c__Gammaproteobacteria|o__Aeromonadales|f__Aeromonadaceae	0.93	0	0	0	0	0	0
k__Bacteria|p__Proteobacteria|c__Gammaproteobacteria|o__Aeromonadales|f__Aeromonadaceae|g__Aeromonas	0.93	0	0	0	0	0	0
k__Bacteria|p__Proteobacteria|c__Gammaproteobacteria|o__Aeromonadales|f__Aeromonadaceae|g__Aeromonas| s__Aeromonas_caviae	0.69	0	0	0	0	0	0
k__Bacteria|p__Proteobacteria|c__Gammaproteobacteria|o__Aeromonadales|f__Aeromonadaceae|g__Aeromonas| s__Aeromonas_caviae|t__Aeromonas_caviae_unclassified	0.69	0	0	0	0	0	0
k__Bacteria|p__Proteobacteria|c__Gammaproteobacteria|o__Aeromonadales|f__Aeromonadaceae|g__Aeromonas| s__Aeromonas_media	0.24	0	0	0	0	0	0
k__Bacteria|p__Proteobacteria|c__Gammaproteobacteria|o__Aeromonadales|f__Aeromonadaceae|g__Aeromonas| s__Aeromonas_media|t_Aeromonas_media_unclassified	0.24	0	0	0	0	0	0

k__Bacteria|p__Proteobacteria|c__Gammaproteobacteria|o__Alteromonadales	1.4	0	0	0	0	0	0
k__Bacteria|p__Proteobacteria|c__Gammaproteobacteria|o__Alteromonadales|f__Alteromonadaceae	0.25	0	0	0	0	0	0
k__Bacteria|p__Proteobacteria|c__Gammaproteobacteria|o__Alteromonadales|f__Alteromonadaceae|g__Alishewanella	0.25	0	0	0	0	0	0
k__Bacteria|p__Proteobacteria|c__Gammaproteobacteria|o__Alteromonadales|f__Alteromonadaceae|g__Alishewanella|s__Alishewanella_agri	0.25	0	0	0	0	0	0
k__Bacteria|p__Proteobacteria|c__Gammaproteobacteria|o__Alteromonadales|f__Alteromonadaceae|g__Alishewanella|s__Alishewanella_agri|t__GCF_000272005	0.25	0	0	0	0	0	0
k__Bacteria|p__Proteobacteria|c__Gammaproteobacteria|o__Alteromonadales|f__Shewanellaceae	1.16	0	0	0	0	0	0
k__Bacteria|p__Proteobacteria|c__Gammaproteobacteria|o__Alteromonadales|f__Shewanellaceae|g__Shewanella	1.16	0	0	0	0	0	0
k__Bacteria|p__Proteobacteria|c__Gammaproteobacteria|o__Alteromonadales|f__Shewanellaceae|g__Shewanella|s__Shewanella_xiamenensis	1.16	0	0	0	0	0	0
k__Bacteria|p__Proteobacteria|c__Gammaproteobacteria|o__Alteromonadales|f__Shewanellaceae|g__Shewanella|s__Shewanella_xiamenensis|t_Shewanella_xiamenensis_unclassified	1.16	0	0	0	0	0	0
k__Bacteria|p__Proteobacteria|c__Gammaproteobacteria|o__Enterobacterales	7.72	41.51	11.48	0.33	0.07	0.25	0.05
k__Bacteria|p__Proteobacteria|c__Gammaproteobacteria|o__Enterobacterales|f__Enterobacteriaceae	7.72	41.51	11.48	0.33	0.07	0.25	0.05
k__Bacteria|p__Proteobacteria|c__Gammaproteobacteria|o__Enterobacteriales|f__Enterobacteriaceae|g__Enterobacter	0	0.07	0	0	0	0	0
k__Bacteria|p__Proteobacteria|c__Gammaproteobacteria|o__Enterobacteriales|f__Enterobacteriaceae|g__Enterobacter|s__Enterobacter_cloacae	0	0.07	0	0	0	0	0
k__Bacteria|p__Proteobacteria|c__Gammaproteobacteria|o__Enterobacteriales|f__Enterobacteriaceae|g__Enterobacter|s__Enterobacter_cloacae|t__Enterobacter_cloacae_unclassified	0	0.07	0	0	0	0	0
k__Bacteria|p__Proteobacteria|c__Gammaproteobacteria|o__Enterobacterales|f__Enterobacteriaceae|g__Escherichia	7.72	41.09	11.39	0.33	0.07	0.25	0.05
k__Bacteria|p__Proteobacteria|c__Gammaproteobacteria|o__Enterobacterales|f__Enterobacteriaceae|g__Escherichia|s__Escherichia_coli	5.28	5.97	6.28	0.33	0.07	0.11	0.05
k__Bacteria|p__Proteobacteria|c__Gammaproteobacteria|o__Enterobacterales|f__Enterobacteriaceae|g__Escherichia|s__Escherichia_coli|t_Escherichia_coli_unclassified	5.28	5.97	6.28	0.33	0.07	0.11	0.05
k__Bacteria|p__Proteobacteria|c__Gammaproteobacteria|o__Enterobacteriales|f__Enterobacteriaceae|g__Escherichia|	2.45	25.12	5.11	0	0	0.14	0

s__Escherichia_unclassified							
k__Bacteria|p__Proteobacteria|c__Gammaproteobacteria|o__Enterobacterales|f__Enterobacteriaceae|g__Klebsiella	0	0.34	0.09	0	0	0	0
k__Bacteria|p__Proteobacteria|c__Gammaproteobacteria|o__Enterobacteriales|f__Enterobacteriaceae|g__Klebsiella|s__Klebsiella_pneumoniae	0	0.16	0	0	0	0	0
k__Bacteria|p__Proteobacteria|c__Gammaproteobacteria|o__Enterobacteriales|f__Enterobacteriaceae|g__Klebsiella|s__Klebsiella_pneumoniae|t__Klebsiella_pneumoniae_unclassified	0	0.16	0	0	0	0	0
k__Bacteria|p__Proteobacteria|c__Gammaproteobacteria|o__Enterobacterales|f__Enterobacteriaceae|g__Klebsiella|s__Klebsiella_variicola	0	0	0.09	0	0	0	0
k__Bacteria|p__Proteobacteria|c__Gammaproteobacteria|o__Enterobacterales|f__Enterobacteriaceae|g__Klebsiella|s__Klebsiella_variicola|t_Klebsiella_variicola_unclassified	0	0	0.09	0	0	0	0
k__Bacteria|p__Proteobacteria|c__Gammaproteobacteria|o__Enterobacteriales|f__Enterobacteriaceae|g__Klebsiella|s__Klebsiella_unclassified	0	0.19	0	0	0	0	0
k__Bacteria|p__Proteobacteria|c__Gammaproteobacteria|o__Oceanospirillales	0	0	0	0	0.29	0	0
k__Bacteria|p__Proteobacteria|c__Gammaproteobacteria|o__Oceanospirillales|f__Halomonadaceae	0	0	0	0	0.29	0	0
k__Bacteria|p__Proteobacteria|c__Gammaproteobacteria|o__Oceanospirillales|f__Halomonadaceae|g__Halomonas	0	0	0	0	0.29	0	0
k__Bacteria|p__Proteobacteria|c__Gammaproteobacteria|o__Oceanospirillales|f__Halomonadaceae|g__Halomonas|s__Halomonas_unclassified	0	0	0	0	0.29	0	0
k__Bacteria|p__Proteobacteria|c__Gammaproteobacteria|o__Pseudomonadales	11.46	1.03	2.27	0.08	0	0	0.21
k__Bacteria|p__Proteobacteria|c__Gammaproteobacteria|o__Pseudomonadales|f__Moraxellaceae	0.87	0.42	0.66	0.08	0	0	0.08
k__Bacteria|p__Proteobacteria|c__Gammaproteobacteria|o__Pseudomonadales|f__Moraxellaceae|g__Acinetobacter	0.87	0.36	0.66	0.08	0	0	0.08
k__Bacteria|p__Proteobacteria|c__Gammaproteobacteria|o__Pseudomonadales|f__Moraxellaceae|g__Acinetobacter|s__Acinetobacter_baumannii	0	0.06	0	0	0	0	0
k__Bacteria|p__Proteobacteria|c__Gammaproteobacteria|o__Pseudomonadales|f__Moraxellaceae|g__Acinetobacter|s__Acinetobacter_baumannii|t__Acinetobacter_baumannii_unclassified	0	0.06	0	0	0	0	0
k__Bacteria|p__Proteobacteria|c__Gammaproteobacteria|o__Pseudomonadales|f__Moraxellaceae|g__Acinetobacter|s__Acinetobacter_idrijaensis	0.23	0	0	0	0	0	0
k__Bacteria|p__Proteobacteria|c__Gammaproteobacteria|o__Pseudomonadales|f__Moraxellaceae|g__Acinetobacter|	0.23	0	0	0	0	0	0

s__Acinetobacter_idrijaensis|t_Acinetobacter_idrijaensis_unclassified							
k__Bacteria|p__Proteobacteria|c__Gammaproteobacteria|o__Pseudomonadales|f__Moraxellaceae|g__Acinetobacter| s__Acinetobacter_johnsonii	0.27	0.02	0	0.07	0	0	0.08
k__Bacteria|p__Proteobacteria|c__Gammaproteobacteria|o__Pseudomonadales|f__Moraxellaceae|g__Acinetobacter| s__Acinetobacter_johnsonii|t__Acinetobacter_johnsonii_unclassified	0.27	0.02	0	0.07	0	0	0.08
k__Bacteria|p__Proteobacteria|c__Gammaproteobacteria|o__Pseudomonadales|f__Moraxellaceae|g__Acinetobacter| s__Acinetobacter_lwoffii	0.09	0.18	0.43	0	0	0	0
k__Bacteria|p__Proteobacteria|c__Gammaproteobacteria|o__Pseudomonadales|f__Moraxellaceae|g__Acinetobacter| s__Acinetobacter_lwoffii|t__Acinetobacter_lwoffii_unclassified	0.09	0.18	0.43	0	0	0	0
k__Bacteria|p__Proteobacteria|c__Gammaproteobacteria|o__Pseudomonadales|f__Moraxellaceae|g__Acinetobacter| s__Acinetobacter_unclassified	0.28	0.1	0.14	0.01	0	0	0
k__Bacteria|p__Proteobacteria|c__Gammaproteobacteria|o__Pseudomonadales|f__Moraxellaceae|g__Enhydrobacter	0	0.06	0.09	0	0	0	0
k__Bacteria|p__Proteobacteria|c__Gammaproteobacteria|o__Pseudomonadales|f__Moraxellaceae|g__Enhydrobacter| s__Enhydrobacter_aerosaccus	0	0.06	0.09	0	0	0	0
k__Bacteria|p__Proteobacteria|c__Gammaproteobacteria|o__Pseudomonadales|f__Moraxellaceae|g__Enhydrobacter| s__Enhydrobacter_aerosaccus|t__GCF_000175915	0	0.06	0.09	0	0	0	0
k__Bacteria|p__Proteobacteria|c__Gammaproteobacteria|o__Pseudomonadales|f__Pseudomonadaceae	10.59	0.6	1.61	0	0	0	0.12
k__Bacteria|p__Proteobacteria|c__Gammaproteobacteria|o__Pseudomonadales|f__Pseudomonadaceae|g__Pseudomonas	10.59	0.6	1.61	0	0	0	0.12
k__Bacteria|p__Proteobacteria|c__Gammaproteobacteria|o__Pseudomonadales|f__Pseudomonadaceae|g__Pseudomonas| s__Pseudomonas_alcaliphila	1.27	0	0	0	0	0	0
k__Bacteria|p__Proteobacteria|c__Gammaproteobacteria|o__Pseudomonadales|f__Pseudomonadaceae|g__Pseudomonas| s__Pseudomonas_alcaliphila|t__Pseudomonas_alcaliphila_unclassified	1.27	0	0	0	0	0	0
k__Bacteria|p__Proteobacteria|c__Gammaproteobacteria|o__Pseudomonadales|f__Pseudomonadaceae|g__Pseudomonas| s__Pseudomonas_fragi	0.03	0.3	0.17	0	0	0	0
k__Bacteria|p__Proteobacteria|c__Gammaproteobacteria|o__Pseudomonadales|f__Pseudomonadaceae|g__Pseudomonas| s__Pseudomonas_fragi|t__Pseudomonas_fragi_unclassified	0.03	0.3	0.17	0	0	0	0
k__Bacteria|p__Proteobacteria|c__Gammaproteobacteria|o__Pseudomonadales|f__Pseudomonadaceae|g__Pseudomonas| s__Pseudomonas_mendocina	0.77	0.01	0	0	0	0	0
k__Bacteria|p__Proteobacteria|c__Gammaproteobacteria|o__Pseudomonadales|f__Pseudomonadaceae|g__Pseudomonas| s__Pseudomonas_mendocina|t__Pseudomonas_mendocina_unclassified	0.77	0.01	0	0	0	0	0

k__Bacteria|p__Proteobacteria|c__Gammaproteobacteria|o__Pseudomonadales|f__Pseudomonadaceae|g__Pseudomonas|s__Pseudomonas_psychrotolerans	0.59	0	0	0	0	0	0
k__Bacteria|p__Proteobacteria|c__Gammaproteobacteria|o__Pseudomonadales|f__Pseudomonadaceae|g__Pseudomonas|s__Pseudomonas_psychrotolerans|t__GCF_000236825	0.59	0	0	0	0	0	0
k__Bacteria|p__Proteobacteria|c__Gammaproteobacteria|o__Pseudomonadales|f__Pseudomonadaceae|g__Pseudomonas|s__Pseudomonas_stutzeri	0.37	0.02	0	0	0	0	0
k__Bacteria|p__Proteobacteria|c__Gammaproteobacteria|o__Pseudomonadales|f__Pseudomonadaceae|g__Pseudomonas|s__Pseudomonas_stutzeri|t__Pseudomonas_stutzeri_unclassified	0.37	0.02	0	0	0	0	0
k__Bacteria|p__Proteobacteria|c__Gammaproteobacteria|o__Pseudomonadales|f__Pseudomonadaceae|g__Pseudomonas|s__Pseudomonas_unclassified	7.56	0.27	1.44	0	0	0	0.12
k__Bacteria|p__Proteobacteria|c__Gammaproteobacteria|o__Xanthomonadales	0.36	0.55	0.28	0	0	0	0
k__Bacteria|p__Proteobacteria|c__Gammaproteobacteria|o__Xanthomonadales|f__Sinobacteraceae	0.36	0.55	0.28	0	0	0	0
k__Bacteria|p__Proteobacteria|c__Gammaproteobacteria|o__Xanthomonadales|f__Sinobacteraceae|g__Sinobacteraceae_unclassified	0.36	0.55	0.28	0	0	0	0
k__Bacteria|p__Proteobacteria|c__Hydrogenophilalia	0.95	0	0	0	0	0	0
k__Bacteria|p__Proteobacteria|c__Hydrogenophilalia|o__Hydrogenophilales	0.95	0	0	0	0	0	0
k__Bacteria|p__Proteobacteria|c__Hydrogenophilalia|o__Hydrogenophilales|f__Hydrogenophilaceae	0.95	0	0	0	0	0	0
k__Bacteria|p__Proteobacteria|c__Hydrogenophilalia|o__Hydrogenophilales|f__Hydrogenophilaceae|g__Hydrogenophilus	0.95	0	0	0	0	0	0
k__Bacteria|p__Proteobacteria|c__Hydrogenophilalia|o__Hydrogenophilales|f__Hydrogenophilaceae|g__Hydrogenophilus|s__Hydrogenophilus_thermoluteolus	0.95	0	0	0	0	0	0
k__Bacteria|p__Proteobacteria|c__Hydrogenophilalia|o__Hydrogenophilales|f__Hydrogenophilaceae|g__Hydrogenophilus|s__Hydrogenophilus_thermoluteolus|t__Hydrogenophilus_thermoluteolus_unclassified	0.95	0	0	0	0	0	0
k__Viruses	12.7	3.65	8.52	0	0.35	0	2.89
k__Viruses|p__Viruses_noname	12.7	3.65	8.52	0	0.35	0	2.89
k__Viruses|p__Viruses_noname|c__Viruses_noname	12.7	3.65	8.52	0	0.35	0	2.89

k__Viruses|p__Viruses_noname|c__Viruses_noname|o__Caudovirales	12.7	3.65	8.52	0	0.35	0	2.89
k__Viruses|p__Viruses_noname|c__Viruses_noname|o__Caudovirales|f__Siphoviridae	12.7	3.65	8.52	0	0.35	0	2.89
k__Viruses|p__Viruses_noname|c__Viruses_noname|o__Caudovirales|f__Siphoviridae|g__C2likevirus	0	0	0.11	0	0	0	0
k__Viruses|p__Viruses_noname|c__Viruses_noname|o__Caudovirales|f__Siphoviridae|g__C2likevirus|s__C2likevirus_unclassified	0	0	0.11	0	0	0	0
k__Viruses|p__Viruses_noname|c__Viruses_noname|o__Caudovirales|f__Siphoviridae|g__Lambdalikevirus	0	0	0	0	0	0	2.89
k__Viruses|p__Viruses_noname|c__Viruses_noname|o__Caudovirales|f__Siphoviridae|g__Lambdalikevirus|s__Enterobacteria_phage_lambda	0	0	0	0	0	0	2.89
k__Viruses|p__Viruses_noname|c__Viruses_noname|o__Caudovirales|f__Siphoviridae|g__Lambdalikevirus|s__Enterobacteria_phage_lambda|t__PRJNA14204	0	0	0	0	0	0	2.89
k__Viruses|p__Viruses_noname|c__Viruses_noname|o__Caudovirales|f__Siphoviridae|g__Siphoviridae_noname	12.57	3.65	8.41	0	0.35	0	0
k__Viruses|p__Viruses_noname|c__Viruses_noname|o__Caudovirales|f__Siphoviridae|g__Siphoviridae_noname|s__Propionibacterium_phage_P100D	0	0	3.46	0	0	0	0
k__Viruses|p__Viruses_noname|c__Viruses_noname|o__Caudovirales|f__Siphoviridae|g__Siphoviridae_noname|s__Propionibacterium_phage_P100D|t__PRJNA177534	0	0	3.46	0	0	0	0
k__Viruses|p__Viruses_noname|c__Viruses_noname|o__Caudovirales|f__Siphoviridae|g__Siphoviridae_noname|s__Propionibacterium_phage_P101A	12.57	3.65	4.96	0	0	0	0
k__Viruses|p__Viruses_noname|c__Viruses_noname|o__Caudovirales|f__Siphoviridae|g__Siphoviridae_noname|s__Propionibacterium_phage_P101A|t__PRJNA177531	12.57	3.65	4.96	0	0	0	0
k__Viruses|p__Viruses_noname|c__Viruses_noname|o__Caudovirales|f__Siphoviridae|g__Siphoviridae_noname|s__Propionibacterium_phage_P104A	0	0	0	0	0.35	0	0
k__Viruses|p__Viruses_noname|c__Viruses_noname|o__Caudovirales|f__Siphoviridae|g__Siphoviridae_noname|s__Propionibacterium_phage_P104A|t__PRJNA177532	0	0	0	0	0.35	0	0
k__Viruses|p__Viruses_noname|c__Viruses_noname|o__Caudovirales|f__Siphoviridae|g__Yualikevirus	0.13	0	0	0	0	0	0
k__Viruses|p__Viruses_noname|c__Viruses_noname|o__Caudovirales|f__Siphoviridae|g__Yualikevirus|s__Yualikevirus_unclassified	0.13	0	0	0	0	0	0

Appendix 3: List of total predicted genes in HsAP

DAL (LTW0001)	GAL (LTW0002)	BLA (LTW0003)
(2E,6E)-farnesyl diphosphate synthase (EC 2.5.1.10)	[Fe] hydrogenase, HymA subunit, putative	(2E,6Z)-farnesyl diphosphate synthase (EC 2.5.1.68)
(R)-citramalate synthase (EC 2.3.1.182)	1-deoxy-D-xylulose 5-phosphate synthase (EC 2.2.1.7)	(R)-citramalate synthase (EC 2.3.1.182)
(S)-3-O-geranylgeranylglyceryl phosphate synthase	1-pyrroline-4-hydroxy-2-carboxylate deaminase (EC 3.5.4.22)	[NiFe] hydrogenase metallocenter assembly protein HypC
[NiFe] hydrogenase metallocenter assembly protein HypC	1,2-dihydroxy-3-keto-5-methylthiopentene dioxygenase (EC 1.13.11.54)	[NiFe] hydrogenase metallocenter assembly protein HypD
[NiFe] hydrogenase metallocenter assembly protein HypD	1,4-alpha-glucan (glycogen) branching enzyme, GH-13-type (EC 2.4.1.18)	[NiFe] hydrogenase metallocenter assembly protein HypE
[NiFe] hydrogenase metallocenter assembly protein HypE	2-aminoethylphosphonate ABC transporter permease protein I (TC 3.A.1.9.1)	[NiFe] hydrogenase metallocenter assembly protein HypF
[NiFe] hydrogenase metallocenter assembly protein HypF	2-dehydro-3-deoxygluconate aldolase (EC 4.1.2-)	[Protein-PII] uridylyltransferase (EC 2.7.7.59)
[NiFe] hydrogenase nickel incorporation protein HypA	2-dehydro-3-deoxyphosphogluconate aldolase (EC 4.1.2.14)	1-acyl-sn-glycerol-3-phosphate acyltransferase (EC 2.3.1.51)
[NiFe] hydrogenase nickel incorporation-associated protein HypB	2-hydroxy-3-keto-5-methylthiopentenyl-1-phosphate phosphatase	1-deoxy-D-xylulose 5-phosphate reductoisomerase (EC 1.1.1.267)
[Protein-PII] uridylyltransferase (EC 2.7.7.59)	2-hydroxy-6-oxo-6-phenylhexa-2,4-dienoate hydrolase (EC 3.7.1.-)	1-deoxy-D-xylulose 5-phosphate synthase (EC 2.2.1.7)
1-acyl-sn-glycerol-3-phosphate acyltransferase (EC 2.3.1.51)	2-keto-3-deoxy-D-arabino-heptulosonate-7-phosphate synthase II (EC 2.5.1.54)	1-hydroxy-2-methyl-2-(E)-butenyl 4-diphosphate synthase (EC 1.17.7.1)
1-deoxy-D-xylulose 5-phosphate reductoisomerase (EC 1.1.1.267)	2-ketogluconate kinase (EC 2.7.1.13)	1-phosphofructokinase (EC 2.7.1.56)
1-deoxy-D-xylulose 5-phosphate synthase (EC 2.2.1.7)	2-ketogluconate transporter	1-pyrroline-4-hydroxy-2-carboxylate deaminase (EC 3.5.4.22)
1-hydroxy-2-methyl-2-(E)-butenyl 4-diphosphate synthase (EC 1.17.7.1)	2-ketogluconate utilization repressor PtxS	1,4-alpha-glucan (glycogen) branching enzyme, GH-13-type (EC 2.4.1.18)
1-phosphofructokinase (EC 2.7.1.56)	2-methylcitrate dehydratase (EC 4.2.1.79)	16S rRNA processing protein RimM
1-pyrroline-4-hydroxy-2-carboxylate deaminase (EC 3.5.4.22)	2-methylcitrate dehydratase FeS dependent (EC 4.2.1.79)	2-amino-3-ketobutyrate coenzyme A ligase (EC 2.3.1.29)
1,4-alpha-glucan (glycogen) branching enzyme, GH-13-type (EC 2.4.1.18)	2-methylcitrate synthase (EC 2.3.3.5)	2-amino-4-hydroxy-6-hydroxymethyldihydropteridine pyrophosphokinase (EC 2.7.6.3)
16S rRNA processing protein RimM	2-methylisocitrate dehydratase (EC 4.2.1.99)	2-aminoethylphosphonate ABC transporter ATP-binding protein (TC 3.A.1.9.1)
2-amino-3-ketobutyrate coenzyme A ligase (EC 2.3.1.29)	2-succinyl-5-enolpyruvyl-6-hydroxy-3-cyclohexene-1-carboxylic-acid synthase (EC 2.2.1.9)	2-aminoethylphosphonate:pyruvate aminotransferase (EC 2.6.1.37)
2-amino-4-hydroxy-6-hydroxymethyldihydropteridine pyrophosphokinase (EC 2.7.6.3)	2-succinyl-6-hydroxy-2,4-cyclohexadiene-1-carboxylate synthase (EC 4.2.99.20)	2-C-methyl-D-erythritol 2,4-cyclodiphosphate synthase (EC 4.6.1.12)
2-aminoethylphosphonate:pyruvate aminotransferase (EC 2.6.1.37)	2,3-bisphosphoglycerate-independent phosphoglycerate mutase, archaeal type (EC 5.4.2.1)	2-C-methyl-D-erythritol 4-phosphate cytidylyltransferase (EC 2.7.7.60)
2-C-methyl-D-erythritol 4-phosphate cytidylyltransferase (EC 2.7.7.60)	2,3-butanediol dehydrogenase, R-alcohol forming, (R)- and (S)-acetoin-specific (EC 1.1.1.4)	2-dehydro-3-deoxygalactonokinase (EC 2.7.1.58)
2-dehydro-3-deoxyglucarate aldolase (EC 4.1.2.20)	2,3-butanediol dehydrogenase, S-alcohol forming, (R)-acetoin-specific (EC	2-dehydro-3-deoxyglucarate aldolase (EC 4.1.2.20)

2-dehydro-3-deoxygluconate kinase (EC 2.7.1.45)
2-dehydro-3-deoxyphosphogalactonate aldolase (EC 4.1.2.21)
2-dehydro-3-deoxyphosphogluconate aldolase (EC 4.1.2.14)
2-dehydropantoate 2-reductase (EC 1.1.1.169)
2-hydroxy-3-oxopropionate reductase (EC 1.1.1.60)
2-hydroxy-6-ketonona-2,4-dienedioic acid hydrolase (EC 3.7.1.-)
2-hydroxy-6-oxo-6-phenylhexa-2,4-dienoate hydrolase (EC 3.7.1.-)
2-hydroxyhepta-2,4-diene-1,7-dioate isomerase (EC 5.3.3.-)
2-isopropylmalate synthase (EC 2.3.3.13)
2-keto-3-deoxy-D-arabino-heptulosonate-7-phosphate synthase I alpha (EC 2.5.1.54)
2-keto-3-deoxy-D-arabino-heptulosonate-7-phosphate synthase II (EC 2.5.1.54)
2-Keto-3-deoxy-D-manno-octulosonate-8-phosphate synthase (EC 2.5.1.55)
2-keto-4-pentenoate hydratase/2-oxohepta-3-ene-1,7-dioic acid hydratase (catechol pathway)
2-Keto-D-gluconate dehydrogenase (EC 1.1.99.4), membrane-bound, flavoprotein
2-ketoaldonate reductase, broad specificity (EC 1.1.1.215) (EC 1.1.1.-)
2-ketogluconate transporter
2-ketoglutaric semialdehyde dehydrogenase (EC 1.2.1.26)
2-methylcitrate dehydratase (EC 4.2.1.79)
2-methylcitrate dehydratase FeS dependent (EC 4.2.1.79)
2-methylcitrate synthase (EC 2.3.3.5)
2-octaprenyl-3-methyl-6-methoxy-1,4-benzoquinol hydroxylase (EC 1.14.13.-)
2-octaprenyl-6-methoxyphenol hydroxylase (EC 1.14.13.-)
2-oxo-hepta-3-ene-1,7-dioic acid hydratase (EC 4.2.-.-)
2-Oxobutyrate oxidase, putative
2-oxoglutarate dehydrogenase E1 component (EC 1.2.4.2)
2-polyprenyl-6-methoxyphenol hydroxylase and related FAD-dependent oxidoreductases

1.1.1.4)
2,3-diketo-5-methylthiopentyl-1-phosphate enolase
2,3-diketo-5-methylthiopentyl-1-phosphate enolase-phosphatase (EC 3.1.3.77)
3-dehydroquinate dehydratase I (EC 4.2.1.10)
3-dehydroquinate synthase (EC 4.2.3.4)
3-hydroxyacyl-CoA dehydrogenase (EC 1.1.1.35)
3-hydroxyanthranilate 3,4-dioxygenase (EC 1.13.11.6)
3-hydroxybutyrate dehydrogenase (EC 1.1.1.30)
3-hydroxybutyryl-CoA dehydratase (EC 4.2.1.55)
3-hydroxybutyryl-CoA dehydrogenase (EC 1.1.1.157)
3-hydroxybutyryl-CoA epimerase (EC 5.1.2.3)
3-keto-L-gulonate 6-phosphate decarboxylase
3-ketoacyl-CoA thiolase [isoleucine degradation] (EC 2.3.1.16)
3-oxoacyl-[acyl-carrier-protein] reductase of FASI (EC 1.1.1.100)
3-oxoacyl-[acyl-carrier-protein] synthase, KASI (EC 2.3.1.41)
3-oxoacyl-[acyl-carrier-protein] synthase, KASI, alternative (EC 2.3.1.41)
3-phenylpropionate dioxygenase, alpha subunit (EC 1.14.12.19)
3-phenylpropionate dioxygenase, beta subunit (EC 1.14.12.19)
3'-to-5' oligoribonuclease (orn)
3'-to-5' oligoribonuclease B, Bacillus type
4-amino, 4-deoxyprephenate dehydrogenase (EC 1.3.1.-)
4-aminobutyraldehyde dehydrogenase (EC 1.2.1.19)
4-hydroxybenzoate transporter
4-hydroxybenzoyl-CoA thioesterase family active site
4'-phosphopantetheinyl transferase (EC 2.7.8.-)
5-amino-6-(5-phosphoribosylamino)uracil reductase (EC 1.1.1.193)
5-aminolevulinate synthase (EC 2.3.1.37)

2-dehydro-3-deoxygluconate kinase (EC 2.7.1.45)
2-dehydro-3-deoxyphosphogluconate aldolase (EC 4.1.2.14)
2-dehydropantoate 2-reductase (EC 1.1.1.169)
2-deoxy-D-gluconate 3-dehydrogenase (EC 1.1.1.125)
2-deoxyglucose-6-phosphate hydrolase YniC
2-hydroxy-3-oxopropionate reductase (EC 1.1.1.60)
2-hydroxy-6-oxo-6-phenylhexa-2,4-dienoate hydrolase (EC 3.7.1.-)
2-hydroxychromene-2-carboxylate isomerase
2-iminoacetate synthase (ThiH) (EC 4.1.99.19)
2-isopropylmalate synthase (EC 2.3.3.13)
2-keto-3-deoxy-D-arabino-heptulosonate-7-phosphate synthase I alpha (EC 2.5.1.54)
2-keto-3-deoxy-D-arabino-heptulosonate-7-phosphate synthase I beta (EC 2.5.1.54)
2-keto-3-deoxy-D-arabino-heptulosonate-7-phosphate synthase II (EC 2.5.1.54)
2-Keto-3-deoxy-D-manno-octulosonate-8-phosphate synthase (EC 2.5.1.55)
2-keto-3-deoxygluconate permease (KDG permease)
2-keto-4-pentenoate hydratase (EC 4.2.1.80)
2-keto-4-pentenoate hydratase/2-oxohepta-3-ene-1,7-dioic acid hydratase (catechol pathway)
2-Keto-D-gluconate dehydrogenase (EC 1.1.99.4), membrane-bound, flavoprotein
2-ketoaldonate reductase, broad specificity (EC 1.1.1.215) (EC 1.1.1.-)
2-ketogluconate 6-phosphate reductase (EC 1.1.1.43)
2-ketogluconate kinase (EC 2.7.1.13)
2-ketogluconate transporter
2-ketoglutaric semialdehyde dehydrogenase (EC 1.2.1.26)
2-methylcitrate dehydratase (EC 4.2.1.79)
2-methylcitrate dehydratase FeS dependent (EC 4.2.1.79)
2-methylcitrate synthase (EC 2.3.3.5)

2-polyprenylphenol hydroxylase and related flavodoxin oxidoreductases
2-succinyl-5-enolpyruvyl-6-hydroxy-3-cyclohexene-1-carboxylic-acid synthase (EC 2.2.1.9)
2-succinyl-6-hydroxy-2,4-cyclohexadiene-1-carboxylate synthase (EC 4.2.99.20)
2,3-bisphosphoglycerate-independent phosphoglycerate mutase (EC 5.4.2.1)
2,3-butanediol dehydrogenase, R-alcohol forming, (R)- and (S)-acetoin-specific (EC 1.1.1.4)
2,3-dihydro-2,3-dihydroxybenzoate dehydrogenase (EC 1.3.1.28) [enterobactin] siderophore
2,3-dihydroxy-2,3-dihydro-phenylpropionate dehydrogenase (EC 1.3.1.-)
2,3-dihydroxybenzoate-AMP ligase (EC 2.7.7.58) [enterobactin] siderophore
2,3-dihydroxybiphenyl 1,2-dioxygenase
2,3,4,5-tetrahydropyridine-2,6-dicarboxylate N-succinyltransferase (EC 2.3.1.117)
2,4-dienoyl-CoA reductase [NADPH] (EC 1.3.1.34)
2,5-diketo-D-gluconic acid reductase B (EC 1.1.1.274)
2'-5' RNA ligase
2',3'-cyclic-nucleotide 2'-phosphodiesterase (EC 3.1.4.16)
3-(3-hydroxy-phenyl)propionate hydroxylase (EC 1.14.13.-)
3-carboxy-cis,cis-muconate cycloisomerase (EC 5.5.1.2)
3-dehydroquinate dehydratase II (EC 4.2.1.10)
3-dehydroquinate synthase (EC 4.2.3.4)
3-demethylubiquinone-9 3-methyltransferase (EC 2.1.1.64)
3-deoxy-D-manno-octulosonate 8-phosphate phosphatase (EC 3.1.3.45)
3-deoxy-manno-octulosonate cytidylyltransferase (EC 2.7.7.38)
3-hydroxyacyl-[acyl-carrier-protein] dehydratase, FabZ form (EC 4.2.1.59)
3-hydroxybutyryl-CoA dehydratase (EC 4.2.1.55)
3-hydroxybutyryl-CoA dehydrogenase (EC 1.1.1.157)
3-hydroxyisobutyrate dehydrogenase (EC 1.1.1.31)

5-Enolpyruvylshikimate-3-phosphate synthase (EC 2.5.1.19)
5-FCL-like protein
5-methylthioribose kinase (EC 2.7.1.100)
5-oxoprolinase (EC 3.5.2.9)
5'-methylthioadenosine nucleosidase (EC 3.2.2.16)
5'-methylthioadenosine/S-adenosylhomocysteine nucleosidase related protein BA2564
6 kDa early secretory antigenic target ESAT-6 (EsxA)
6-Methyl salicylic acid synthase
6-phosphogluconolactonase (EC 3.1.1.31)
6-phosphogluconolactonase (EC 3.1.1.31), eukaryotic type
6-pyruvoyl tetrahydrobiopterin synthase (EC 4.2.3.12)
ABC transporter involved in cytochrome c biogenesis, ATPase component CcmA
ABC-type Fe3+-siderophore transport system, permease 2 component
ABC-type nitrate/sulfonate/bicarbonate transport system, ATPase component
Acetaldehyde dehydrogenase (acetylating) (EC 1.2.1.10)
Acetaldehyde dehydrogenase, acetylating, (EC 1.2.1.10) in gene cluster for degradation of phenols, cresols, catechol
Acetaldehyde dehydrogenase, ethanolamine utilization cluster
Acetate kinase (EC 2.7.2.1)
Acetoacetate metabolism regulatory protein AtoC
Acetoacetyl-CoA reductase (EC 1.1.1.36)
Acetoacetyl-CoA synthetase (EC 6.2.1.16)
Acetoacetyl-CoA synthetase [leucine] (EC 6.2.1.16)
Acetoin (diacetyl) reductase (EC 1.1.1.5)
Acetoin catabolism protein X
Acetoin dehydrogenase E1 component alpha-subunit (EC 1.2.4.-)

2-octaprenyl-3-methyl-6-methoxy-1,4-benzoquinol hydroxylase (EC 1.14.13.-)
2-octaprenyl-6-methoxyphenol hydroxylase (EC 1.14.13.-)
2-oxo-hepta-3-ene-1,7-dioic acid hydratase (EC 4.2.-.-)
2-Oxobutyrate oxidase, putative
2-oxoglutarate dehydrogenase E1 component (EC 1.2.4.2)
2-polyprenyl-6-methoxyphenol hydroxylase and related FAD-dependent oxidoreductases
2-polyprenylphenol hydroxylase and related flavodoxin oxidoreductases
2-succinyl-5-enolpyruvyl-6-hydroxy-3-cyclohexene-1-carboxylic-acid synthase (EC 2.2.1.9)
2,3-bisphosphoglycerate-independent phosphoglycerate mutase (EC 5.4.2.1)
2,3-butanediol dehydrogenase, R-alcohol forming, (R)- and (S)-acetoin-specific (EC 1.1.1.4)
2,3-dihydro-2,3-dihydroxybenzoate dehydrogenase (EC 1.3.1.28) [enterobactin] siderophore
2,3-dihydroxybenzoate-AMP ligase (EC 2.7.7.58) [enterobactin] siderophore
2,3,4,5-tetrahydropyridine-2,6-dicarboxylate N-acetyltransferase (EC 2.3.1.89)
2,3,4,5-tetrahydropyridine-2,6-dicarboxylate N-succinyltransferase (EC 2.3.1.117)
2,4-dienoyl-CoA reductase [NADPH] (EC 1.3.1.34)
2,5-diketo-D-gluconate reductase A (EC 1.1.1.274)
2,5-diketo-D-gluconic acid reductase B (EC 1.1.1.274)
2'-5' RNA ligase
2',3'-cyclic-nucleotide 2'-phosphodiesterase (EC 3.1.4.16)
3-(3-hydroxy-phenyl)propionate hydroxylase (EC 1.14.13.-)
3-carboxy-cis,cis-muconate cycloisomerase (EC 5.5.1.2)
3-dehydro-L-gulonate 2-dehydrogenase (EC 1.1.1.130)
3-dehydroquinate dehydratase I (EC 4.2.1.10)
3-dehydroquinate dehydratase II (EC 4.2.1.10)
3-dehydroquinate synthase (EC 4.2.3.4)

3-hydroxyisobutyryl-CoA hydrolase (EC 3.1.2.4)	Acetolactate synthase small subunit (EC 2.2.1.6), predicted	3-demethylubiquinone-9 3-methyltransferase (EC 2.1.1.64)
3-isopropylmalate dehydratase large subunit (EC 4.2.1.33)	Acetolactate synthase small subunit (EC 2.2.1.6), Xanthomonadales type	3-deoxy-D-manno-octulosonate 8-phosphate phosphatase (EC 3.1.3.45)
3-isopropylmalate dehydratase small subunit (EC 4.2.1.33)	Acetyl-CoA acetyltransferase (EC 2.3.1.9)	3-deoxy-manno-octulosonate cytidylyltransferase (EC 2.7.7.38)
3-isopropylmalate dehydrogenase (EC 1.1.1.85)	Acetyl-CoA:acetoacetyl-CoA transferase, alpha subunit (EC 2.8.3.8)	3-hydroxyacyl-[acyl-carrier-protein] dehydratase, FabZ form (EC 4.2.1.59)
3-ketoacyl-CoA thiolase (EC 2.3.1.16)	Acetyl-coenzyme A carboxyl transferase alpha chain (EC 6.4.1.2)	3-hydroxyacyl-CoA dehydrogenase (EC 1.1.1.35)
3-methyl-2-oxobutanoate hydroxymethyltransferase (EC 2.1.2.11)	Acetyl-coenzyme A carboxyl transferase beta chain (EC 6.4.1.2)	3-hydroxybutyryl-CoA dehydratase (EC 4.2.1.55)
3-methylmercaptopropionyl-CoA dehydrogenase (DmdC)	Actin-assembly inducing protein ActA precursor	3-hydroxybutyryl-CoA dehydrogenase (EC 1.1.1.157)
3-methylmercaptopropionyl-CoA ligase (DmdB)	Activator of the mannose operon (transcriptional antiterminator), BglG family	3-hydroxyisobutyrate dehydrogenase (EC 1.1.1.31)
3-oxoacyl-[acyl-carrier protein] reductase (EC 1.1.1.100)	Acyl carrier protein associated with anthrachelin biosynthesis	3-hydroxyisobutyryl-CoA hydrolase (EC 3.1.2.4)
3-oxoacyl-[acyl-carrier-protein] synthase, KASI (EC 2.3.1.41)	Acyl-CoA dehydrogenase, short-chain specific (EC 1.3.99.2)	3-isopropylmalate dehydratase large subunit (EC 4.2.1.33)
3-oxoadipate CoA-transferase subunit A (EC 2.8.3.6)	Acylphosphate phosphohydrolase (EC 3.6.1.7), putative	3-isopropylmalate dehydratase small subunit (EC 4.2.1.33)
3-oxoadipate CoA-transferase subunit B (EC 2.8.3.6)	Additional component NikL of nickel ECF transporter	3-isopropylmalate dehydrogenase (EC 1.1.1.85)
3-phenylpropionate dioxygenase, alpha subunit (EC 1.14.12.19)	Additional substrate-specific component CbiN of cobalt ECF transporter	3-ketoacyl-CoA thiolase (EC 2.3.1.16)
3-polyprenyl-4-hydroxybenzoate carboxy-lyase (EC 4.1.1.-)	Aerobactin siderophore receptor IutA	3-methyl-2-oxobutanoate hydroxymethyltransferase (EC 2.1.2.11)
3,4-dihydroxy-2-butanone 4-phosphate synthase (EC 4.1.99.12)	Aerobic glycerol-3-phosphate dehydrogenase (EC 1.1.5.3)	3-methylmercaptopropionyl-CoA dehydrogenase (DmdC)
3,4-dihydroxyphenylacetate 2,3-dioxygenase (EC 1.13.11.15)	Agmatinase (EC 3.5.3.11)	3-methylmercaptopropionyl-CoA ligase (DmdB)
3'-to-5' exoribonuclease RNase R	Agmatine deiminase (EC 3.5.3.12)	3-oxoacyl-[acyl-carrier protein] reductase (EC 1.1.1.100)
3'-to-5' oligoribonuclease (orn)	Agmatine/putrescine antiporter, associated with agmatine catabolism	3-oxoacyl-[acyl-carrier-protein] synthase, KASI (EC 2.3.1.41)
3',5'-cyclic-nucleotide phosphodiesterase (EC 3.1.4.17)	Alanine-anticapsin ligase BacD	3-oxoadipate CoA-transferase subunit A (EC 2.8.3.6)
4-alpha-glucanotransferase (amylomaltase) (EC 2.4.1.25)	Alcohol dehydrogenase (EC 1.1.1.1)	3-oxoadipate CoA-transferase subunit B (EC 2.8.3.6)
4-aminobutyraldehyde dehydrogenase (EC 1.2.1.19)	Aldehyde dehydrogenase A (EC 1.2.1.22)	3-phenylpropionate dioxygenase ferredoxin subunit
4-carboxymuconolactone decarboxylase (EC 4.1.1.44)	Aldehyde dehydrogenase B (EC 1.2.1.22)	3-phenylpropionate dioxygenase, beta subunit (EC 1.14.12.19)
4-cresol dehydrogenase [hydroxylating] cytochrome c subunit precursor	Alfa-L-rhamnosidase (EC 3.2.1.40)	3-polyprenyl-4-hydroxybenzoate carboxy-lyase (EC 4.1.1.-)
4-cresol dehydrogenase [hydroxylating] flavoprotein subunit (EC 1.17.99.1)	Alkan-1-ol dehydrogenase, PQQ-dependent (EC 1.1.99.20)	3-polyprenyl-4-hydroxybenzoate carboxy-lyase UbiX (EC 4.1.1.-)
4-diphosphocytidyl-2-C-methyl-D-erythritol kinase (EC 2.7.1.148)	Alkanesulfonate monooxygenase (EC 1.14.14.5)	3,4-dihydroxy-2-butanone 4-phosphate synthase (EC 4.1.99.12)
4-hydroxy-3-methylbut-2-enyl diphosphate reductase (EC 1.17.1.2)	Alkanesulfonates ABC transporter ATP-binding protein	3,4-dihydroxyphenylacetate 2,3-dioxygenase (EC 1.13.11.15)
4-hydroxybenzoate transporter	Allophycocyanin alpha chain	3'-to-5' exoribonuclease RNase R
4-hydroxybenzoyl-CoA reductase, beta subunit (EC 1.3.99.20)	Allophycocyanin beta chain	3'-to-5' oligoribonuclease (orn)

4-hydroxybenzoyl-CoA reductase, gamma subunit (EC 1.3.99.20)
4-hydroxyphenylpyruvate dioxygenase (EC 1.13.11.27)
4-hydroxyproline epimerase (EC 5.1.1.8)
4-hydroxythreonine-4-phosphate dehydrogenase (EC 1.1.1.262)
4-keto-6-deoxy-N-Acetyl-D-hexosaminyl-(Lipid carrier) aminotransferase
4'-phosphopantetheinyl transferase (EC 2.7.8.-) [enterobactin] siderophore
4Fe-4S ferredoxin, nitrogenase-associated
5-amino-6-(5-phosphoribosylamino)uracil reductase (EC 1.1.1.193)
5-aminolevulinate synthase (EC 2.3.1.37)
5-carboxymethyl-2-hydroxymuconate semialdehyde dehydrogenase (EC 1.2.1.60)
5-carboxymethyl-2-oxo-hex-3- ene-1,7-dioate decarboxylase (EC 4.1.1.68)
5-dehydro-4-deoxyglucarate dehydratase (EC 4.2.1.41)
5-Enolpyruvylshikimate-3-phosphate synthase (EC 2.5.1.19)
5-FCL-like protein
5-keto-D-gluconate 5-reductase (EC 1.1.1.69)
5-methyl-dCTP pyrophosphohydrolase (EC 3.6.1.-)
5-methyltetrahydrofolate--homocysteine methyltransferase (EC 2.1.1.13)
5-methyltetrahydropteroyltriglutamate--homocysteine methyltransferase (EC 2.1.1.14)
5-methylthioribose kinase (EC 2.7.1.100)
5-nucleotidase SurE (EC 3.1.3.5)
5-oxoprolinase (EC 3.5.2.9)
5,10-methylenetetrahydrofolate reductase (EC 1.5.1.20)
5'-methylthioadenosine nucleosidase (EC 3.2.2.16)
5'-methylthioadenosine phosphorylase (EC 2.4.2.28)
5'-nucleotidase (EC 3.1.3.5)
5'-nucleotidase family protein in cluster with NagD-like phosphatase
5'-nucleotidase YjjG (EC 3.1.3.5)
6-phospho-beta-glucosidase (EC 3.2.1.86)

Allophycocyanin-B
Alpha-1,2-mannosidase
Alpha-acetolactate decarboxylase (EC 4.1.1.5)
Alpha-glucosidase, family 31 of glycosyl hydrolases, COG1501
Alternative L-arabinose isomerase (EC 5.3.1.4)
Altronate dehydratase (EC 4.2.1.7)
Amino acid-binding ACT
Aminoglycoside N6'-acetyltransferase (EC 2.3.1.82)
Anaerobic nitric oxide reductase flavorubredoxin
Antiphagocytic M protein
Arabinose operon protein AraL
Arabinose-regulated TonB-dependent outer membrane receptor
Archaeosine cluster gene X
AreB (Aryl-alcohol dehydrogenase) (EC 1.1.1.90)
Arginine decarboxylase (EC 4.1.1.19)
Arginine decarboxylase, catabolic (EC 4.1.1.19)
Arginine deiminase (EC 3.5.3.6)
Arginine/agmatine antiporter
Arginine/ornithine antiporter
Arginine/ornithine antiporter ArcD
Ascorbate-specific PTS system, EIIB component (EC 2.7.1.69)
Asp-X dipeptidase
Asparaginyl-tRNA synthetase (EC 6.1.1.22)
Aspartate aminotransferase (EC 2.6.1.1)
Aspartyl-tRNA(Asn) amidotransferase subunit A (EC 6.3.5.6)
Aspartyl-tRNA(Asn) amidotransferase subunit C (EC 6.3.5.6)
Aspartyl-tRNA(Asn) synthetase (EC 6.1.1.23)
ATP-dependent nuclease, subunit B

3',5'-cyclic-nucleotide phosphodiesterase (EC 3.1.4.17)
3'(2'),5'-bisphosphate nucleotidase (EC 3.1.3.7)
4-alpha-glucanotransferase (amylomaltase) (EC 2.4.1.25)
4-aminobutyraldehyde dehydrogenase (EC 1.2.1.19)
4-carboxymuconolactone decarboxylase (EC 4.1.1.44)
4-cresol dehydrogenase [hydroxylating] flavoprotein subunit (EC 1.17.99.1)
4-deoxy-L-threo-5-hexosulose-uronate ketol-isomerase (EC 5.3.1.17)
4-diphosphocytidyl-2-C-methyl-D-erythritol kinase (EC 2.7.1.148)
4-hydroxy-3-methylbut-2-enyl diphosphate reductase (EC 1.17.1.2)
4-hydroxybenzoate transporter
4-hydroxybenzoyl-CoA reductase, gamma subunit (EC 1.3.99.20)
4-hydroxybenzoyl-CoA thioesterase family active site
4-hydroxyphenylacetate 3-monooxygenase (EC 1.14.13.3)
4-hydroxyphenylpyruvate dioxygenase (EC 1.13.11.27)
4-hydroxythreonine-4-phosphate dehydrogenase (EC 1.1.1.262)
4-keto-6-deoxy-N-Acetyl-D-hexosaminyl-(Lipid carrier) aminotransferase
4-oxalocrotonate decarboxylase (EC 4.1.1.77)
4'-phosphopantetheinyl transferase (EC 2.7.8.-)
4Fe-4S ferredoxin, nitrogenase-associated
5-amino-6-(5-phosphoribosylamino)uracil reductase (EC 1.1.1.193)
5-aminolevulinate synthase (EC 2.3.1.37)
5-carboxymethyl-2-hydroxymuconate semialdehyde dehydrogenase (EC 1.2.1.60)
5-carboxymethyl-2-oxo-hex-3- ene-1,7-dioate decarboxylase (EC 4.1.1.68)
5-dehydro-4-deoxyglucarate dehydratase (EC 4.2.1.41)
5-deoxy-glucuronate isomerase (EC 5.3.1.-)
5-Enolpyruvylshikimate-3-phosphate synthase (EC 2.5.1.19)
5-FCL-like protein
5-formyltetrahydrofolate cyclo-ligase (EC 6.3.3.2)

6-phosphofructokinase (EC 2.7.1.11)
6-phosphofructokinase class II (EC 2.7.1.11)
6-phosphogluconate dehydrogenase, decarboxylating (EC 1.1.1.44)
6-phosphogluconolactonase (EC 3.1.1.31), eukaryotic type
6-pyruvoyl tetrahydrobiopterin synthase (EC 4.2.3.12)
6,7-dimethyl-8-ribityllumazine synthase (EC 2.5.1.78)
8-amino-7-oxononanoate synthase (EC 2.3.1.47)
A/G-specific adenine glycosylase (EC 3.2.2.-)
ABC Fe3+ siderophore transporter, inner membrane subunit
ABC transporter involved in cytochrome c biogenesis, ATPase component CcmA
ABC transporter, periplasmic spermidine putrescine-binding protein PotD (TC 3.A.1.11.1)
ABC transporter, permease protein
ABC-type nitrate/sulfonate/bicarbonate transport system, ATPase component
ABC-type nitrate/sulfonate/bicarbonate transport system, permease component
ABC-type nitrate/sulfonate/bicarbonate transport systems, periplasmic components
ABC-type sugar transport system, periplasmic binding protein YcjN
ABC-type tungstate transport system, periplasmic binding protein
Acetaldehyde dehydrogenase (EC 1.2.1.10)
Acetaldehyde dehydrogenase, acetylating, (EC 1.2.1.10) in gene cluster for degradation of phenols, cresols, catechol
Acetate kinase (EC 2.7.2.1)
Acetate permease ActP (cation/acetate symporter)
Acetoacetate decarboxylase (EC 4.1.1.4)
Acetoacetyl-CoA reductase (EC 1.1.1.36)
Acetoacetyl-CoA synthetase (EC 6.2.1.16)
Acetolactate synthase large subunit (EC 2.2.1.6)
Acetolactate synthase small subunit (EC 2.2.1.6)
Acetolactate synthase, catabolic (EC 2.2.1.6)

ATP-dependent RNA helicase PA3950
ATP-dependent RNA helicase SO1501
ATP:Cob(I)alamin adenosyltransferase (EC 2.5.1.17), ethanolamine utilization
ATPase component NikO of energizing module of nickel ECF transporter
Autoinducer 2 (AI-2) kinase LsrK (EC 2.7.1.-)
Bacilysin biosynthesis protein BacB
Bacteriocin-like peptide J BlpJ
Bacterioferritin
Bactriocin immunity protein BlpX
Benzoate transport, extracellular ligand-binding receptor
Beta-galactosidase (EC 3.2.1.23)
Beta-ketoadipyl CoA thiolase (EC 2.3.1.9)
Beta-mannosidase (EC 3.2.1.25)
Beta-ureidopropionase (EC 3.5.1.6)
Beta-xylosidase (EC 3.2.1.37)
Biofilm operon icaABCD HTH-type negative transcriptional regulator IcaR
Biopolymer transport ExbD protein
Biotin carboxyl carrier protein of acetyl-CoA carboxylase
Biotin carboxylase of acetyl-CoA carboxylase (EC 6.3.4.14)
Biotin carboxylase of methylcrotonyl-CoA carboxylase (EC 6.3.4.14)
Biphenyl dioxygenase beta subunit (EC 1.14.12.18)
Biphenyl-2,3-diol 1,2-dioxygenase (EC 1.13.11.39)
biphenyl-2,3-diol 1,2-dioxygenase III-related protein
Branched-chain acyl-CoA dehydrogenase (EC 1.3.99.12)
Branched-chain alpha-keto acid dehydrogenase, E1 component, alpha subunit (EC 1.2.4.4)
Branched-chain amino acid aminotransferase (EC 2.6.1.42)
Broad-specificity glycerol dehydrogenase (EC 1.1.99.22), subunit SldA

5-keto-2-deoxy-D-gluconate-6 phosphate aldolase [form 2] (EC 4.1.2.29)
5-keto-2-deoxygluconokinase (EC 2.7.1.92)
5-keto-D-gluconate 5-reductase (EC 1.1.1.69)
5-methyl-dCTP pyrophosphohydrolase (EC 3.6.1.-)
5-methyltetrahydrofolate--homocysteine methyltransferase (EC 2.1.1.13)
5-methyltetrahydropteroyltriglutamate--homocysteine methyltransferase (EC 2.1.1.14)
5-nucleotidase SurE (EC 3.1.3.5)
5-oxopent-3-ene-1,2,5-tricarboxylate decarboxylase.
5-oxoprolinase (EC 3.5.2.9)
5,10-methylenetetrahydrofolate reductase (EC 1.5.1.20)
5'-methylthioadenosine nucleosidase (EC 3.2.2.16)
5'-methylthioadenosine phosphorylase (EC 2.4.2.28)
5'-nucleotidase (EC 3.1.3.5)
5'-nucleotidase YjjG (EC 3.1.3.5)
6-phospho-beta-glucosidase (EC 3.2.1.86)
6-phospho-beta-glucosidase ascB (EC 3.2.1.86)
6-phosphofructokinase (EC 2.7.1.11)
6-phosphofructokinase class II (EC 2.7.1.11)
6-phosphogluconate dehydrogenase, decarboxylating (EC 1.1.1.44)
6-phosphogluconolactonase (EC 3.1.1.31)
6-phosphogluconolactonase (EC 3.1.1.31), eukaryotic type
6,7-dimethyl-8-ribityllumazine synthase (EC 2.5.1.78)
8-amino-7-oxononanoate synthase (EC 2.3.1.47)
A/G-specific adenine glycosylase (EC 3.2.2.-)
ABC transporter involved in cytochrome c biogenesis, ATPase component CcmA
ABC transporter involved in cytochrome c biogenesis, CcmB subunit
ABC transporter, periplasmic spermidine putrescine-binding protein PotD (TC 3.A.1.11.1)

Acetyl-CoA acetyltransferase (EC 2.3.1.9)
Acetyl-CoA synthetase (ADP-forming) alpha and beta chains, putative
Acetyl-CoA synthetase (ADP-forming) alpha chain (EC 6.2.1.13)
Acetyl-CoA:acetoacetyl-CoA transferase, alpha subunit (EC 2.8.3.8)
Acetyl-coenzyme A carboxyl transferase alpha chain (EC 6.4.1.2)
Acetyl-coenzyme A carboxyl transferase beta chain (EC 6.4.1.2)
Acetyl-coenzyme A synthetase (EC 6.2.1.1)
Acetylglutamate kinase (EC 2.7.2.8)
Acetylornithine aminotransferase (EC 2.6.1.11)
Acetylornithine deacetylase (EC 3.5.1.16)
Aconitate hydratase (EC 4.2.1.3)
Aconitate hydratase 2 (EC 4.2.1.3)
Acriflavin resistance protein
Acryloyl-CoA reductase AcuI/YhdH (EC 1.3.1.84)
Acting phosphoribosylanthranilate isomerase (EC 5.3.1.24)
Acyclic terpenes utilization regulator AtuR, TetR family
Acyl-[acyl-carrier-protein]--UDP-N-acetylglucosamine O-acyltransferase (EC 2.3.1.129)
Acyl-CoA dehydrogenase, short-chain specific (EC 1.3.99.2)
Acyl-CoA thioesterase II (EC 3.1.2.-)
Acyl-phosphate:glycerol-3-phosphate O-acyltransferase PlsY
Adenine deaminase (EC 3.5.4.2)
Adenine phosphoribosyltransferase (EC 2.4.2.7)
Adenosine (5')-pentaphospho-(5'')-adenosine pyrophosphohydrolase (EC 3.6.1.-)
Adenosine deaminase (EC 3.5.4.4)
Adenosylcobinamide-phosphate guanylyltransferase (EC 2.7.7.62)
Adenosylhomocysteinase (EC 3.3.1.1)
Adenosylmethionine-8-amino-7-oxononanoate aminotransferase (EC 2.6.1.62)
Adenylate cyclase (EC 4.6.1.1)
Broad-substrate range phospholipase C (EC 3.1.4.3)
Butyrate kinase (EC 2.7.2.7)
Butyrate-acetoacetate CoA-transferase subunit B (EC 2.8.3.9)
Butyryl-CoA dehydrogenase (EC 1.3.99.2)
C5a peptidase (EC 3.4.21.-)
Cannabinoid receptor 1
Carbamate kinase (EC 2.7.2.2)
Carbon dioxide concentrating mechanism protein CcmL
Carbon dioxide concentrating mechanism protein CcmN
Carbon monoxide dehydrogenase large chain (EC 1.2.99.2)
Carbon-monoxide dehydrogenase form II, large subunit (EC 1.2.99.2)
Carbonic anhydrase (EC 4.2.1.1)
Carboxynorspermidine decarboxylase, putative (EC 4.1.1.-)
Carboxynorspermidine dehydrogenase, putative (EC 1.1.1.-)
Catalase (EC 1.11.1.6)
Cation efflux system protein CusF precursor
CDP-diacylglycerol--serine O-phosphatidyltransferase (EC 2.7.8.8)
Cell division initiation protein DivIVA
Cell division protein FtsI [Peptidoglycan synthetase] (EC 2.4.1.129)
Cell surface protein IsdA, transfers heme from hemoglobin to apo-IsdC
Cell surface protein Shp, transfers heme from hemoglobin to apo-SiaA/HtsA
Chaperone protein DnaK
Chemotaxis protein CheC -- inhibitor of MCP methylation
Chemotaxis protein CheD
Chemotaxis protein CheV (EC 2.7.3.-)
Chemotaxis protein CheX
Chemotaxis protein methyltransferase CheR (EC 2.1.1.80)
Chemotaxis regulator - transmits chemoreceptor signals to flagelllar motor
ABC-type Fe3+-siderophore transport system, permease 2 component
ABC-type nitrate/sulfonate/bicarbonate transport system, ATPase component
ABC-type nitrate/sulfonate/bicarbonate transport system, permease component
ABC-type nitrate/sulfonate/bicarbonate transport systems, periplasmic components
ABC-type sugar transport system, periplasmic binding protein YcjN
ABC-type tungstate transport system, periplasmic binding protein
Acetaldehyde dehydrogenase, acetylating, (EC 1.2.1.10) in gene cluster for degradation of phenols, cresols, catechol
Acetaldehyde dehydrogenase, ethanolamine utilization cluster
Acetate kinase (EC 2.7.2.1)
Acetate permease ActP (cation/acetate symporter)
Acetoacetate decarboxylase (EC 4.1.1.4)
Acetoacetyl-CoA reductase (EC 1.1.1.36)
Acetoacetyl-CoA synthetase (EC 6.2.1.16)
Acetoin catabolism protein X
Acetoin dehydrogenase E1 component alpha-subunit (EC 1.2.4.-)
Acetoin dehydrogenase E1 component beta-subunit (EC 1.2.4.-)
Acetolactate synthase large subunit (EC 2.2.1.6)
Acetolactate synthase small subunit (EC 2.2.1.6)
Acetolactate synthase, catabolic (EC 2.2.1.6)
Acetyl-CoA acetyltransferase
Acetyl-CoA acetyltransferase (EC 2.3.1.9)
Acetyl-CoA synthetase (ADP-forming) alpha and beta chains, putative
Acetyl-CoA synthetase (ADP-forming) alpha chain (EC 6.2.1.13)
Acetyl-CoA:acetoacetyl-CoA transferase, alpha subunit (EC 2.8.3.8)
Acetyl-CoA:acetoacetyl-CoA transferase, beta subunit (EC 2.8.3.8)
Acetyl-coenzyme A carboxyl transferase alpha chain (EC 6.4.1.2)
Acetyl-coenzyme A carboxyl transferase beta chain (EC 6.4.1.2)
Acetyl-coenzyme A synthetase (EC 6.2.1.1)

Adenylate kinase (EC 2.7.4.3)
Adenylosuccinate lyase (EC 4.3.2.2)
Adenylosuccinate synthetase (EC 6.3.4.4)
Adenylylsulfate kinase (EC 2.7.1.25)
Adhesin of unknown specificity SdrE, similar to bone sialoprotein-binding protein Bbp
ADP compounds hydrolase NudE (EC 3.6.1.-)
ADP-heptose--lipooligosaccharide heptosyltransferase II (EC 2.4.1.-)
ADP-L-glycero-D-manno-heptose-6-epimerase (EC 5.1.3.20)
ADP-ribose pyrophosphatase (EC 3.6.1.13)
Aerobic C4-dicarboxylate transporter for fumarate, L-malate, D-malate, succunate
Aerobic cobaltochelatase CobS subunit (EC 6.6.1.2)
Aerobic cobaltochelatase CobT subunit (EC 6.6.1.2)
Aerobic glycerol-3-phosphate dehydrogenase (EC 1.1.5.3)
Aerotaxis sensor receptor protein
Agmatinase (EC 3.5.3.11)
Alanine dehydrogenase (EC 1.4.1.1)
Alanine racemase (EC 5.1.1.1)
Alanyl-tRNA synthetase (EC 6.1.1.7)
Alcohol dehydrogenase (EC 1.1.1.1)
Alcohol dehydrogenase, zinc-containing
Aldehyde dehydrogenase A (EC 1.2.1.22)
Aldehyde dehydrogenase B (EC 1.2.1.22)
Aldose 1-epimerase (EC 5.1.3.3)
Alkaline phosphatase (EC 3.1.3.1)
Alkanesulfonate monooxygenase (EC 1.14.14.5)
Alkanesulfonate utilization operon LysR-family regulator CbI
Alkanesulfonates ABC transporter ATP-binding protein
Alkanesulfonates transport system permease protein

components CheY
Chemotaxis response - phosphatase CheZ
Chemotaxis response regulator protein-glutamate methylesterase CheB (EC 3.1.1.61)
Chorismate synthase (EC 4.2.3.5)
Co-activator of prophage gene expression IbrA
Co-activator of prophage gene expression IbrB
CoB--CoM heterodisulfide reductase subunit C (EC 1.8.98.1)
Cob(I)alamin adenosyltransferase (EC 2.5.1.17)
Cobalt-zinc-cadmium resistance protein CzcD
Coenzyme B synthesis from 2-oxoglutarate: steps 1, 6, and 10
Colanic acid biosynthesis acetyltransferase WcaF (EC 2.3.1.-)
Colanic acid biosynthesis glycosyl transferase WcaA
Colanic acid biosynthesis glycosyl transferase WcaL
Colanic acid biosynthsis UDP-glucose lipid carrier transferase WcaJ
Colanic acid biosysnthesis protein WcaK
Colicin I receptor precursor
Conjugative signal peptidase TrhF
Conjugative transfer entry exclusion protein TrbK
Conjugative transfer protein ELI_00880
Conjugative transfer protein PSLT087
Conjugative transfer protein s043
Conjugative transfer protein TrbA
Conjugative transfer protein TrbB
Conjugative transfer protein TrbC
Conjugative transfer protein TrbE
Conjugative transfer protein TrbF
Conjugative transfer protein TrbG
Conjugative transfer protein TrbH
Conjugative transfer protein TrbI

Acetylglutamate kinase (EC 2.7.2.8)
Acetylornithine aminotransferase (EC 2.6.1.11)
Acetylornithine deacetylase (EC 3.5.1.16)
Acetylornithine deacetylase/Succinyl-diaminopimelate desuccinylase and related deacylases
Aconitate hydratase (EC 4.2.1.3)
Aconitate hydratase 2 (EC 4.2.1.3)
Acriflavin resistance protein
Acryloyl-CoA reductase AcuI/YhdH (EC 1.3.1.84)
Acyclic terpenes utilization regulator AtuR, TetR family
Acyl carrier protein
Acyl-[acyl-carrier-protein]--UDP-N-acetylglucosamine O-acyltransferase (EC 2.3.1.129)
Acyl-CoA dehydrogenase, short-chain specific (EC 1.3.99.2)
Acyl-CoA thioesterase II (EC 3.1.2.-)
Acyl-phosphate:glycerol-3-phosphate O-acyltransferase PlsY
Acylamino-acid-releasing enzyme (EC 3.4.19.1)
ADA regulatory protein
Adenine deaminase (EC 3.5.4.2)
Adenine phosphoribosyltransferase (EC 2.4.2.7)
Adenosine (5')-pentaphospho-(5'')-adenosine pyrophosphohydrolase (EC 3.6.1.-)
Adenosine deaminase (EC 3.5.4.4)
Adenosylcobinamide-phosphate guanylyltransferase (EC 2.7.7.62)
Adenosylhomocysteinase (EC 3.3.1.1)
Adenosylmethionine-8-amino-7-oxononanoate aminotransferase (EC 2.6.1.62)
Adenylate cyclase (EC 4.6.1.1)
Adenylate kinase (EC 2.7.4.3)
Adenylosuccinate lyase (EC 4.3.2.2)
Adenylosuccinate synthetase (EC 6.3.4.4)
Adenylylsulfate kinase (EC 2.7.1.25)

Alkanesulfonates-binding protein
Alkyl hydroperoxide reductase protein C (EC 1.6.4.-)
Alkyl hydroperoxide reductase protein F (EC 1.6.4.-)
Alkyl hydroperoxide reductase subunit C-like protein
Allantoicase (EC 3.5.3.4)
Allantoin permease
Allantoinase (EC 3.5.2.5)
Allophanate hydrolase (EC 3.5.1.54)
Alpha-amylase (EC 3.2.1.1)
Alpha-galactosidase (EC 3.2.1.22)
Alpha-glucosidase (EC 3.2.1.20)
Alpha-ketoglutarate-dependent taurine dioxygenase (EC 1.14.11.17)
Alpha-L-fucosidase (EC 3.2.1.51)
Alpha-mannosidase (EC 3.2.1.24)
Alpha-N-arabinofuranosidase (EC 3.2.1.55)
Alpha-N-arabinofuranosidase 2 (EC 3.2.1.55)
Alpha-xylosidase (EC 3.2.1.-)
Alpha,alpha-trehalose-phosphate synthase [UDP-forming] (EC 2.4.1.15)
Altronate dehydratase (EC 4.2.1.7)
Altronate oxidoreductase (EC 1.1.1.58)
Amidases related to nicotinamidase
Amidohydrolase YlmB, involved in salvage of thiamin pyrimidine moiety
Amidophosphoribosyltransferase (EC 2.4.2.14)
Amine oxidase [flavin-containing] (EC 1.4.3.4)
Aminoacyl-histidine dipeptidase (Peptidase D) (EC 3.4.13.3)
Aminobenzoyl-glutamate transport protein
Aminodeoxychorismate lyase (EC 4.1.3.38)
Aminomethyltransferase (glycine cleavage system T protein) (EC 2.1.2.10)
Aminopeptidase YpdF (MP-, MA-, MS-, AP-, NP- specific)
Ammonium transporter
AMP nucleosidase (EC 3.2.2.4)

Conjugative transfer protein TrbJ
Conjugative transfer protein TrbL
Conjugative transfer protein TrbO
Conjugative transfer protein TrbP (IncF TraX homolog)
Conjugative transposon protein TraB
Conjugative transposon protein TraD
Conjugative transposon protein TraE
Conjugative transposon protein TraF
Conjugative transposon protein TraG
Conjugative transposon protein TraH
Conjugative transposon protein TraI
Conjugative transposon protein TraJ
Conjugative transposon protein TraK
Conjugative transposon protein TraM
Conjugative transposon protein TraN
Conjugative transposon protein TraP
CooT
Copper homeostasis protein CutE
Copper homeostasis protein CutF precursor
Coproporphyrinogen III oxidase, aerobic (EC 1.3.3.3)
CpeR homolog, phycoerythrin linker-proteins region
CRISPR-associated protein Cas7
Curli production assembly/transport component CsgE
Curli production assembly/transport component CsgG
Cyanobacteria-specific RpoD-like sigma factor, type-7
Cyanophycin synthase II
Cyanophycinase (EC 3.4.15.6)
Cytochrome b6-f complex subunit VI (PetL)
Cytochrome b6-f complex subunit VII (PetM)
Cytochrome b6-f complex subunit VIII (PetN)
Cytochrome b6-f complex subunit, cytochrome b6, putative

Adenylylsulfate reductase alpha-subunit (EC 1.8.99.2)
ADP compounds hydrolase NudE (EC 3.6.1.-)
ADP-heptose--lipooligosaccharide heptosyltransferase II (EC 2.4.1.-)
ADP-L-glycero-D-manno-heptose-6-epimerase (EC 5.1.3.20)
ADP-ribose pyrophosphatase (EC 3.6.1.13)
Aerobic C4-dicarboxylate transporter for fumarate, L-malate, D-malate, succunate
Aerobic cobaltochelatase CobS subunit (EC 6.6.1.2)
Aerobic cobaltochelatase CobT subunit (EC 6.6.1.2)
Aerobic glycerol-3-phosphate dehydrogenase (EC 1.1.5.3)
Aerotaxis sensor receptor protein
Agmatinase (EC 3.5.3.11)
Agmatine deiminase (EC 3.5.3.12)
Alanine dehydrogenase (EC 1.4.1.1)
Alanine racemase (EC 5.1.1.1)
Alanine racemase, biosynthetic (EC 5.1.1.1)
Alanine racemase, catabolic (EC 5.1.1.1)
Alanyl-tRNA synthetase (EC 6.1.1.7)
Alanyl-tRNA synthetase family protein
Alcohol dehydrogenase (EC 1.1.1.1)
Aldehyde dehydrogenase (EC 1.2.1.3)
Aldehyde dehydrogenase B (EC 1.2.1.22)
Aldose 1-epimerase (EC 5.1.3.3)
Aldose 1-epimerase family protein YphB
Alkaline phosphatase (EC 3.1.3.1)
Alkanesulfonate monooxygenase (EC 1.14.14.5)
Alkanesulfonate utilization operon LysR-family regulator CbI
Alkanesulfonates ABC transporter ATP-binding protein
Alkanesulfonates transport system permease protein
Alkanesulfonates-binding protein
Alkyl hydroperoxide reductase protein C (EC 1.6.4.-)
Alkyl hydroperoxide reductase protein F (EC 1.6.4.-)

Anaerobic dehydrogenases, typically selenocysteine-containing
Anaerobic dimethyl sulfoxide reductase chain B (EC 1.8.5.3)
Anaerobic glycerol-3-phosphate dehydrogenase subunit C (EC 1.1.5.3)
Anaerobic nitric oxide reductase flavorubredoxin
Anthranilate dioxygenase reductase
Anthranilate phosphoribosyltransferase (EC 2.4.2.18)
Anthranilate synthase, amidotransferase component (EC 4.1.3.27)
Anthranilate synthase, aminase component (EC 4.1.3.27)
Apolipoprotein N-acyltransferase (EC 2.3.1.-)
Aquaporin Z
Arabinan endo-1,5-alpha-L-arabinosidase (EC 3.2.1.99)
Arabinose 5-phosphate isomerase (EC 5.3.1.13)
Arabinose-proton symporter
Arginase (EC 3.5.3.1)
Arginine ABC transporter, periplasmic arginine-binding protein ArtJ
Arginine ABC transporter, permease protein ArtM
Arginine ABC transporter, permease protein ArtQ
Arginine deiminase (EC 3.5.3.6)
Arginine N-succinyltransferase (EC 2.3.1.109)
Arginine N-succinyltransferase, alpha subunit (EC 2.3.1.109)
Arginine N-succinyltransferase, beta subunit (EC 2.3.1.109)
Arginine pathway regulatory protein ArgR, repressor of arg regulon
Arginine/agmatine antiporter
Arginine/ornithine antiporter ArcD
Argininosuccinate lyase (EC 4.3.2.1)
Argininosuccinate synthase (EC 6.3.4.5)
Arginyl-tRNA synthetase (EC 6.1.1.19)
Arogenate dehydrogenase (EC 1.3.1.43)

Cytochrome c-type biogenesis protein CcmD, interacts with CcmCE
Cytochrome c-type biogenesis protein Ccs1/ResB
Cytochrome c-type biogenesis protein, archaeal, distantly related to heme lyase subunit CcmF
Cytochrome d ubiquinol oxidase subunit II (EC 1.10.3.-)
cytochrome oxidase,subunit I (cydA-1)
cytochrome-c3 hydrogenase delta chain
cytochrome-c3 hydrogenase gamma chain
Cytolysin and hemolysin, HlyA, Pore-forming toxin
Cytoplasmic copper homeostasis protein cutC
D-amino acid dehydrogenase (EC 1.4.99.1) family protein in hydroxy-L-proline catabolic cluster
D-galactonate transporter
D-glycerate 2-kinase (EC 2.7.1.-)
D-mannose isomerase (EC 5.3.1.7)
D-proline reductase, 23 kDa subunit (EC 1.21.4.1)
D-proline reductase, 26 kDa subunit (EC 1.21.4.1)
D-xylose transport ATP-binding protein XylG
Delta(3)-cis-delta(2)-trans-enoyl-CoA isomerase (EC 5.3.3.8)
Deoxyribonuclease TatD
Diaminohydroxyphosphoribosylaminopyrimidine deaminase (EC 3.5.4.26)
Dihydroflavonol-4-reductase (EC 1.1.1.219)
Dihydrolipoamide acetyltransferase component of pyruvate dehydrogenase complex (EC 2.3.1.12)
Dihydrolipoamide dehydrogenase (EC 1.8.1.4)
Dihydrolipoamide dehydrogenase of acetoin dehydrogenase (EC 1.8.1.4)
Dimethylallyltransferase (EC 2.5.1.1)
Dimethylamine methyltransferase corrinoid protein
Dimethylamine:corrinoid methyltransferase
DinG family ATP-dependent helicase YoaA
Dipeptide-binding ABC transporter, periplasmic substrate-binding component (TC 3.A.1.5.2)

Alkyl hydroperoxide reductase subunit C-like protein
Alkylated DNA repair protein AlkB
Alkylphosphonate utilization operon protein PhnA
Allantoate amidohydrolase (EC 3.5.3.9)
Allantoicase (EC 3.5.3.4)
Allantoin permease
Allantoinase (EC 3.5.2.5)
Allophanate hydrolase (EC 3.5.1.54)
Alpha-amylase (EC 3.2.1.1)
Alpha-aspartyl dipeptidase Peptidase E (EC 3.4.13.21)
Alpha-galactosidase (EC 3.2.1.22)
Alpha-glucosidase (EC 3.2.1.20)
Alpha-ketoglutarate-dependent taurine dioxygenase (EC 1.14.11.17)
Alpha-L-fucosidase (EC 3.2.1.51)
Alpha-L-Rha alpha-1,3-L-rhamnosyltransferase (EC 2.4.1.-)
Alpha-mannosidase (EC 3.2.1.24)
Alpha-N-arabinofuranosidase (EC 3.2.1.55)
Alpha-ribazole-5'-phosphate phosphatase (EC 3.1.3.73)
Alpha-xylosidase (EC 3.2.1.-)
Alpha,alpha-trehalose-phosphate synthase [UDP-forming] (EC 2.4.1.15)
Alternative cytochrome c oxidase polypeptide CoxM (EC 1.9.3.1)
Alternative cytochrome c oxidase polypeptide CoxN (EC 1.9.3.1)
Alternative cytochrome c oxidase polypeptide CoxP (EC 1.9.3.1)
Alternative dihydrofolate reductase 3
Altronate dehydratase (EC 4.2.1.7)
Altronate oxidoreductase (EC 1.1.1.58)
Amidase clustered with urea ABC transporter and nitrile hydratase functions
Amidophosphoribosyltransferase (EC 2.4.2.14)

Aromatic amino acid transport protein AroP
Aromatic hydrocarbon utilization transcriptional regulator CatR (LysR family)
Aromatic-L-amino-acid decarboxylase (EC 4.1.1.28)
Arsenate reductase (EC 1.20.4.1)
Arsenic efflux pump protein
Arsenic resistance protein ArsH
Arsenical pump-driving ATPase (EC 3.6.3.16)
Arsenical-resistance protein ACR3
Arylsulfatase (EC 3.1.6.1)
Ascorbate utilization transcriptional regulator UlaR, HTH-type
Ascorbate-specific PTS system, EIIC component
Asp-X dipeptidase
Asparagine synthetase [glutamine-hydrolyzing] (EC 6.3.5.4)
Asparaginyl-tRNA synthetase (EC 6.1.1.22)
Aspartate aminotransferase (EC 2.6.1.1)
Aspartate ammonia-lyase (EC 4.3.1.1)
Aspartate carbamoyltransferase (EC 2.1.3.2)
Aspartate racemase (EC 5.1.1.13)
Aspartate--ammonia ligase (EC 6.3.1.1)
Aspartate-semialdehyde dehydrogenase (EC 1.2.1.11)
Aspartokinase (EC 2.7.2.4)
Aspartyl-tRNA synthetase (EC 6.1.1.12)
Aspartyl-tRNA(Asn) amidotransferase subunit A (EC 6.3.5.6)
Aspartyl-tRNA(Asn) amidotransferase subunit B (EC 6.3.5.6)
Assimilatory nitrate reductase large subunit (EC:1.7.99.4)
ATP phosphoribosyltransferase (EC 2.4.2.17)
ATP phosphoribosyltransferase regulatory subunit (EC 2.4.2.17)
ATP synthase alpha chain (EC 3.6.3.14)

DNA repair and recombination eukaryotic
DNA topoisomerase III, Bacillus plasimid type (EC 5.99.1.2)
DNA-binding domain of ModE
DNA-binding HTH domain in riboflavin kinase
DNA-binding response regulator, LuxR family, near polyamine transporter
DNA-directed RNA polymerase beta' subunit (EC 2.7.7.6)
DNA-directed RNA polymerase delta (= beta") subunit (EC 2.7.7.6)
DNA-directed RNA polymerase delta subunit (EC 2.7.7.6)
DNA-directed RNA polymerase III 17 kDa polypeptide (EC 2.7.7.6)
DNA-directed RNA polymerase III 25 kDa polypeptide (EC 2.7.7.6)
DNA-directed RNA polymerase III 36 kDa polypeptide (EC 2.7.7.6)
DNA-directed RNA polymerase III 47 kDa polypeptide (EC 2.7.7.6)
DNA-directed RNA polymerase III second largest subunit (EC 2.7.7.6)
DNA-directed RNA polymerases I and III 16 kDa polypeptide (EC 2.7.7.6)
DNA-directed RNA polymerases I, II, and III 15 kDa polypeptide (EC 2.7.7.6)
DNA-directed RNA polymerases I, II, and III 7.7 kDa polypeptide (EC 2.7.7.6)
DNA-directed RNA polymerases I, II, and III 8.3 kDa polypeptide (EC 2.7.7.6)
DUF1794
Efflux pump Lde
Electron transfer flavoprotein-ubiquinone oxidoreductase (EC 1.5.5.1)
Electron transfer flavoprotein, alpha subunit
Electron transfer flavoprotein, beta subunit
Endo-alpha-mannosidase
Energy conserving hydrogenase Eha associated protein, ribokinase homolog (protein T)
Energy conserving hydrogenase Ehb protein H
Energy-conserving hydrogenase (ferredoxin), subunit A
Energy-conserving hydrogenase (ferredoxin), subunit B
Energy-conserving hydrogenase (ferredoxin), subunit E

Amine oxidase [flavin-containing] (EC 1.4.3.4)
Aminoacyl-histidine dipeptidase (Peptidase D) (EC 3.4.13.3)
Aminobenzoyl-glutamate transport protein
Aminodeoxychorismate lyase (EC 4.1.3.38)
Aminomethyltransferase (glycine cleavage system T protein) (EC 2.1.2.10)
Aminopeptidase YpdF (MP-, MA-, MS-, AP-, NP- specific)
Ammonium transporter
AMP nucleosidase (EC 3.2.2.4)
Anaerobic dehydrogenases, typically selenocysteine-containing
Anaerobic dimethyl sulfoxide reductase chain B (EC 1.8.5.3)
Anaerobic glycerol-3-phosphate dehydrogenase subunit A (EC 1.1.5.3)
Anaerobic glycerol-3-phosphate dehydrogenase subunit B (EC 1.1.5.3)
Anaerobic glycerol-3-phosphate dehydrogenase subunit C (EC 1.1.5.3)
Anaerobic nitric oxide reductase flavorubredoxin
Anthranilate dioxygenase reductase
Anthranilate phosphoribosyltransferase (EC 2.4.2.18)
Anthranilate synthase, aminase component (EC 4.1.3.27)
Antiholin-like protein LrgA
Apolipoprotein N-acyltransferase (EC 2.3.1.-)
Aquaporin Z
Arabinose 5-phosphate isomerase (EC 5.3.1.13)
Arabinose operon regulatory protein
archaeal ATPase, fused to C-terminal DUF234 domain
Archaeal S-adenosylmethionine synthetase (EC 2.5.1.6)
Arginase (EC 3.5.3.1)
Arginine ABC transporter, permease protein ArtQ
Arginine decarboxylase, catabolic (EC 4.1.1.19)
Arginine deiminase (EC 3.5.3.6)

ATP synthase beta chain (EC 3.6.3.14)
ATP synthase delta chain (EC 3.6.3.14)
ATP synthase epsilon chain (EC 3.6.3.14)
ATP synthase F0 sector subunit a
ATP synthase gamma chain (EC 3.6.3.14)
ATP synthase protein I
ATP-dependent Clp protease adaptor protein ClpS
ATP-dependent Clp protease ATP-binding subunit ClpA
ATP-dependent Clp protease ATP-binding subunit ClpX
ATP-dependent Clp protease proteolytic subunit (EC 3.4.21.92)
ATP-dependent Clp protease, ATP-binding subunit ClpC
ATP-dependent DNA helicase pcrA (EC 3.6.1.-)
ATP-dependent DNA helicase RecG (EC 3.6.1.-)
ATP-dependent DNA helicase RecQ
ATP-dependent DNA helicase Rep
ATP-dependent DNA helicase UvrD/PcrA
ATP-dependent DNA helicase UvrD/PcrA, proteobacterial paralog
ATP-dependent DNA ligase (EC 6.5.1.1) clustered with Ku protein, LigD
ATP-dependent helicase DinG/Rad3
ATP-dependent hsl protease ATP-binding subunit HslU
ATP-dependent nuclease, subunit A
ATP-dependent nuclease, subunit B
ATP-dependent protease HslV (EC 3.4.25.-)
ATP-dependent protease La (EC 3.4.21.53)
ATP-dependent protease La (EC 3.4.21.53) Type I
ATP-dependent protease La (EC 3.4.21.53) Type II
ATP-dependent RNA helicase Atu1833
ATP-dependent RNA helicase DbpA
ATP-dependent RNA helicase NGO0650
ATP-dependent RNA helicase PA3950

Enoyl-[acyl-carrier-protein] reductase [NADH] (EC 1.3.1.9)
Enoyl-CoA hydratase (EC 4.2.1.17)
Enoyl-CoA hydratase [isoleucine degradation] (EC 4.2.1.17)
enoyl-CoA hydratase, R-specific
Enterobactin receptor IrgA
Epimerase KguE
Ethanolamine ammonia-lyase heavy chain (EC 4.3.1.7)
Ethanolamine operon regulatory protein
Ethanolamine sensory transduction histidine kinase
Ethanolamine two-component response regulator
Ethanolamine utilization polyhedral-body-like protein EutL
Ethanolamine utilization polyhedral-body-like protein EutM
Ethanolamine utilization polyhedral-body-like protein EutN
Ethanolamine utilization protein EutQ
Ethanolamine utilization protein similar to PduA/PduJ
Ethanolamine utilization protein similar to PduL
Ethanolamine utilization protein similar to PduU
Evolved beta-D-galactosidase transcriptional repressor
Evolved beta-D-galactosidase, alpha subunit
Evolved beta-D-galactosidase, beta subunit
Exoribonuclease II (EC 3.1.13.1)
Export ABC transporter permease
F420-dependent methylenetetrahydromethanopterin dehydrogenase (EC 1.5.99.9)
FAD-dependent monooxygenase PhzS
Fatty-acid amide hydrolase (EC 3.1.-.-)
Ferredoxin 3 fused to uncharacterized domain
Ferredoxin-dependent glutamate synthase (EC 1.4.7.1)
Ferredoxin-like protein
Ferredoxin-like protein FixX
Ferredoxin-like protein YdiT

Arginine N-succinyltransferase (EC 2.3.1.109)
Arginine pathway regulatory protein ArgR, repressor of arg regulon
Arginine utilization regulatory protein RocR
Arginine-tRNA-protein transferase (EC 2.3.2.8)
Arginine/agmatine antiporter
Arginine/ornithine antiporter ArcD
Argininosuccinate lyase (EC 4.3.2.1)
Argininosuccinate synthase (EC 6.3.4.5)
Arginyl-tRNA synthetase (EC 6.1.1.19)
Aromatic amino acid transport protein AroP
Aromatic hydrocarbon utilization transcriptional regulator CatR (LysR family)
Aromatic-L-amino-acid decarboxylase (EC 4.1.1.28)
Arsenate reductase (EC 1.20.4.1)
Arsenic efflux pump protein
Arsenic resistance protein ArsH
Arsenical-resistance protein ACR3
Arylsulfatase (EC 3.1.6.1)
AscBF operon repressor
Ascorbate utilization transcriptional regulator UlaR, HTH-type
Ascorbate-specific PTS system, EIIB component (EC 2.7.1.69)
Ascorbate-specific PTS system, EIIC component
Asp-X dipeptidase
Asparagine synthetase [glutamine-hydrolyzing] (EC 6.3.5.4)
Asparaginyl-tRNA synthetase (EC 6.1.1.22)
Aspartate aminotransferase (EC 2.6.1.1)
Aspartate ammonia-lyase (EC 4.3.1.1)
Aspartate carbamoyltransferase (EC 2.1.3.2)
Aspartate racemase (EC 5.1.1.13)
Aspartate--ammonia ligase (EC 6.3.1.1)
Aspartate-semialdehyde dehydrogenase (EC 1.2.1.11)

ATP-dependent RNA helicase RhlB
ATP-dependent RNA helicase RhlE
ATP-dependent RNA helicase SrmB
ATP-dependent RNA helicase VCA0768
ATPase component BioM of energizing module of biotin ECF transporter
ATPase component CbiO of energizing module of cobalt ECF transporter
ATPase provides energy for both assembly of type IV secretion complex and secretion of T-DNA complex (VirB11)
ATPase provides energy for both assembly of type IV secretion complex and secretion of T-DNA complex (VirB4)
ATPase, AFG1 family
Autoinducer 2 (AI-2) ABC transport system, membrane channel protein LsrD
Autoinducer 2 (AI-2) aldolase LsrF (EC 4.2.1.-)
Autoinducer 2 (AI-2) modifying protein LsrG
Bacitracin export ATP-binding protein BceA
Bacitracin export permease protein BceB
Benzoate 1,2-dioxygenase (EC 1.14.12.10)
benzoate degradation ring-cleavage hydrolase
benzoate MFS transporter BenK
Benzoate transport protein
Benzoate transport related protein
Benzoate transport, ATPase component
Benzoate transport, extracellular ligand-binding receptor
Benzoate transport, inner-membrane translocator
Benzoate transport, inner-membrane translocator precursor
Benzoate-CoA ligase (EC 6.2.1.25)
Benzoyl-CoA oxygenase component A
Benzoyl-CoA oxygenase component B

Ferredoxin-like protein YgcO
Ferredoxin-NADP(+) reductase (EC 1.18.1.2)
Ferric aerobactin ABC transporter, periplasmic substrate binding protein
Ferric aerobactin ABC transporter, permease component
Ferric hydroxamate ABC transporter (TC 3.A.1.14.3), periplasmic substrate binding protein FhuD
Ferric receptor CfrA
Ferric siderophore transport system, biopolymer transport protein ExbD
Ferric siderophore transport system, periplasmic binding protein TonB
Ferric vibriobactin, enterobactin transport system, permease protein ViuG (TC 3.A.1.14.6)
Ferric vibriobactin, enterobactin transport system, substrate-binding protein ViuP (TC 3.A.1.14.6)
Ferrichrome-iron receptor
Ferrochelatase, protoheme ferro-lyase (EC 4.99.1.1)
FIG003492: Threonine dehydrogenase and related Zn-dependent dehydrogenases
FIG009886: phosphoesterase
FIG031860: hypothetical protein, Rv2337c
FIG146085: 3'-to-5' oligoribonuclease A, Bacillus type
Flagellar biosynthesis protein FlhB
Flagellar motor switch protein FliG
Flagellar motor switch protein FliM
Flavodoxin
Flavodoxin 2
Flavohemoprotein (Hemoglobin-like protein) (Flavohemoglobin) (Nitric oxide dioxygenase) (EC 1.14.12.17)
FMN adenylyltransferase (EC 2.7.7.2)
Formiminotetrahydrofolate cyclodeaminase (EC 4.3.1.4)
Formylmethanofuran dehydrogenase (molybdenum) operon gene G
Formylmethanofuran dehydrogenase (molybdenum) subunit C (EC 1.2.99.5)

Aspartokinase (EC 2.7.2.4)
Aspartyl-tRNA synthetase (EC 6.1.1.12)
Aspartyl-tRNA(Asn) amidotransferase subunit A (EC 6.3.5.6)
Aspartyl-tRNA(Asn) amidotransferase subunit B (EC 6.3.5.6)
Assimilatory nitrate reductase large subunit (EC:1.7.99.4)
ATP phosphoribosyltransferase (EC 2.4.2.17)
ATP phosphoribosyltransferase regulatory subunit (EC 2.4.2.17)
ATP synthase alpha chain (EC 3.6.3.14)
ATP synthase beta chain (EC 3.6.3.14)
ATP synthase delta chain (EC 3.6.3.14)
ATP synthase epsilon chain (EC 3.6.3.14)
ATP synthase F0 sector subunit a
ATP synthase F0 sector subunit c
ATP synthase gamma chain (EC 3.6.3.14)
ATP-dependent Clp protease ATP-binding subunit ClpA
ATP-dependent Clp protease ATP-binding subunit ClpX
ATP-dependent Clp protease proteolytic subunit (EC 3.4.21.92)
ATP-dependent Clp protease, ATP-binding subunit ClpC
ATP-dependent DNA helicase pcrA (EC 3.6.1.-)
ATP-dependent DNA helicase RecG (EC 3.6.1.-)
ATP-dependent DNA helicase RecQ
ATP-dependent DNA helicase Rep
ATP-dependent DNA ligase (EC 6.5.1.1)
ATP-dependent DNA ligase (EC 6.5.1.1) clustered with Ku protein, LigD
ATP-dependent DNA ligase (EC 6.5.1.1) LigC
ATP-dependent hsl protease ATP-binding subunit HslU

Benzoylformate decarboxylase (EC 4.1.1.7)
Beta N-acetyl-glucosaminidase (EC 3.2.1.52)
Beta-carotene ketolase (EC 1.14.-.-)
Beta-fimbriae chaperone protein
Beta-galactosidase (EC 3.2.1.23)
Beta-glucosidase (EC 3.2.1.21)
Beta-glucuronidase (EC 3.2.1.31)
Beta-hexosaminidase (EC 3.2.1.52)
Beta-ketoadipate enol-lactone hydrolase (EC 3.1.1.24)
Beta-ketoadipyl CoA thiolase (EC 2.3.1.-)
Beta-lactamase (EC 3.5.2.6)
Beta-lactamase class C and other penicillin binding proteins
Beta-lactamase class D
Beta-mannosidase (EC 3.2.1.25)
Beta-ureidopropionase (EC 3.5.1.6)
Betaine aldehyde dehydrogenase (EC 1.2.1.8)
Biopolymer transport protein ExbD/TolR
Biosynthetic arginine decarboxylase (EC 4.1.1.19)
Biosynthetic Aromatic amino acid aminotransferase alpha (EC 2.6.1.57)
Biosynthetic Aromatic amino acid aminotransferase beta (EC 2.6.1.57)
Biotin carboxyl carrier protein of acetyl-CoA carboxylase
Biotin carboxylase of acetyl-CoA carboxylase (EC 6.3.4.14)
Biotin synthase (EC 2.8.1.6)
Biphenyl-2,3-diol 1,2-dioxygenase (EC 1.13.11.39)
biphenyl-2,3-diol 1,2-dioxygenase III-related protein
Bis(5'-nucleosyl)-tetraphosphatase (asymmetrical) (EC 3.6.1.17)
Blue copper oxidase CueO precursor

Formylmethanofuran dehydrogenase (tungsten) subunit D (EC 1.2.99.5)
Formylmethanofuran dehydrogenase subunit B (EC 1.2.99.5)
Formylmethanofuran--tetrahydromethanopterin N-formyltransferase (EC 2.3.1.101)
Fosfomycin resistance protein FosB
Fosfomycin resistance protein FosX
Fructokinase in mannoside utilization gene cluster (EC 2.7.1.4)
Fructose ABC transporter, substrate-binding component FrcB
Galactose-regulated TonB-dependent outer membrane receptor
Galactoside O-acetyltransferase (EC 2.3.1.18)
Gene SCO4494, often clustered with other genes in menaquinone via futalosine pathway
Glucitol operon repressor
Glucitol/sorbitol-specific transport protein GutA
Glucoamylase (EC 3.2.1.3)
Glucokinase (EC 2.7.1.2)
Gluconate 2-dehydrogenase (EC 1.1.99.3), membrane-bound, gamma subunit
Gluconate dehydratase (EC 4.2.1.39)
Glucose dehydrogenase (EC 1.1.99.10), membrane-bound, flavoprotein
Glucose-6-phosphate 1-dehydrogenase (EC 1.1.1.49)
Glutamate racemase (EC 5.1.1.3)
Glutamate synthase [NADPH] large chain (EC 1.4.1.13)
Glutamate synthase [NADPH] putative GlxC chain (EC 1.4.1.13)
Glutamate-1-semialdehyde aminotransferase (EC 5.4.3.8)
Glutamate-ammonia-ligase adenylyltransferase (EC 2.7.7.42)
Glutamine amidotransferase, class-II
Glutamine synthetase inactivating factor IF17
Glutamine synthetase type II, eukaryotic (EC 6.3.1.2)
Glutamine synthetase, clostridia type (EC 6.3.1.2)

ATP-dependent protease HslV (EC 3.4.25.-)
ATP-dependent protease La (EC 3.4.21.53)
ATP-dependent protease La (EC 3.4.21.53) Type I
ATP-dependent protease La (EC 3.4.21.53) Type II
ATP-dependent RNA helicase Atu1833
ATP-dependent RNA helicase Bcep18194_A5658
ATP-dependent RNA helicase DbpA
ATP-dependent RNA helicase NGO0650
ATP-dependent RNA helicase RhlB
ATP-dependent RNA helicase RhlE
ATP-dependent RNA helicase SrmB
ATP:Cob(I)alamin adenosyltransferase (EC 2.5.1.17), ethanolamine utilization
ATPase component CbiO of energizing module of cobalt ECF transporter
ATPase provides energy for both assembly of type IV secretion complex and secretion of T-DNA complex (VirB11)
ATPase provides energy for both assembly of type IV secretion complex and secretion of T-DNA complex (VirB4)
ATPase, AFG1 family
Autoinducer 2 (AI-2) ABC transport system, fused AI2 transporter subunits and ATP-binding component
Autoinducer 2 (AI-2) ABC transport system, membrane channel protein LsrC
Autoinducer 2 (AI-2) ABC transport system, membrane channel protein LsrD
Autoinducer 2 (AI-2) ABC transport system, periplasmic AI-2 binding protein LsrB
Autoinducer 2 (AI-2) aldolase LsrF (EC 4.2.1.-)
Autoinducer 2 (AI-2) kinase LsrK (EC 2.7.1.-)
Autolysis histidine kinase LytS
Bacterioferritin
Benzoate 1,2-dioxygenase (EC 1.14.12.10)
Benzoate 1,2-dioxygenase alpha subunit (EC 1.14.12.10)
benzoate degradation ring-cleavage hydrolase

Branched-chain acyl-CoA dehydrogenase (EC 1.3.99.12)
Branched-chain alpha-keto acid dehydrogenase, E1 component, alpha subunit (EC 1.2.4.4)
Branched-chain alpha-keto acid dehydrogenase, E1 component, beta subunit (EC 1.2.4.4)
Branched-chain amino acid ABC transporter, amino acid-binding protein (TC 3.A.1.4.1)
Branched-chain amino acid aminotransferase (EC 2.6.1.42)
Branched-chain amino acid transport ATP-binding protein LivF (TC 3.A.1.4.1)
Branched-chain amino acid transport ATP-binding protein LivG (TC 3.A.1.4.1)
Branched-chain amino acid transport system permease protein LivM (TC 3.A.1.4.1)
Butyryl-CoA dehydrogenase (EC 1.3.99.2)
C-terminal domain of CinA type S
cAMP-binding proteins - catabolite gene activator and regulatory subunit of cAMP-dependent protein kinases
cAMP-dependent Kef-type K+ transport system
Carbamate kinase (EC 2.7.2.2)
Carbamoyl-phosphate synthase large chain (EC 6.3.5.5)
Carbamoyl-phosphate synthase small chain (EC 6.3.5.5)
carbon monoxide dehydrogenase D protein
carbon monoxide dehydrogenase E protein
carbon monoxide dehydrogenase G protein
Carbon monoxide dehydrogenase large chain (EC 1.2.99.2)
Carbon monoxide dehydrogenase large chain (EC 1.2.99.2) parolog without usual motifs
Carbon monoxide dehydrogenase medium chain (EC 1.2.99.2)
Carbon monoxide dehydrogenase small chain (EC 1.2.99.2)
Carbon starvation protein A
Carbon-monoxide dehydrogenase form II, large subunit (EC 1.2.99.2)
Carbonic anhydrase (EC 4.2.1.1)
Carboxynorspermidine dehydrogenase, putative (EC 1.1.1.-)
Cardiolipin synthetase (EC 2.7.8.-)
Glutaminyl-tRNA synthetase (EC 6.1.1.18)
Glutamyl-tRNA reductase (EC 1.2.1.70)
Glutamyl-tRNA synthetase (EC 6.1.1.17)
Glutamyl-tRNA(Gln) amidotransferase asparaginase subunit (EC 6.3.5.7)
Glutamyl-tRNA(Gln) amidotransferase subunit A (EC 6.3.5.7)
Glutamyl-tRNA(Gln) amidotransferase subunit B (EC 6.3.5.7)
Glutamyl-tRNA(Gln) amidotransferase subunit C (EC 6.3.5.7)
Glutamyl-tRNA(Gln) amidotransferase subunit F (EC 6.3.5.7)
Glutamyl-tRNA(Gln) amidotransferase transferase subunit (EC 6.3.5.7)
Glutamyl-tRNA(Gln) synthetase (EC 6.1.1.24)
Glutaredoxin 1
Glutaredoxin 2
Glutaredoxin 3 (Grx3)
Glutathione peroxidase (EC 1.11.1.9)
Glutathione S-transferase, unnamed subgroup 2 (EC 2.5.1.18)
Glutathione-regulated potassium-efflux system ATP-binding protein
Glycerol dehydratase reactivation factor small subunit
Glycerol dehydratase small subunit (EC 4.2.1.30)
Glycine N-methyltransferase (EC 2.1.1.20)
Glycine riboswitch
Glycolate oxidase (EC 1.1.3.15)
Glycosyltransferase PglI (EC 2.4.1.-)
Glycosyltransferase SypQ
GTP cyclohydrolase II (EC 3.5.4.25)
GTP cyclohydrolase III (EC 3.5.4.29)
Haemin uptake system outer membrane receptor
Hcp transcriptional regulator HcpR (Crp/Fnr family)
benzoate MFS transporter BenK
Benzoate transport, ATPase component
Benzoate transport, extracellular ligand-binding receptor
Benzoate transport, inner-membrane translocator
Benzoate transport, inner-membrane translocator precursor
Benzoate-CoA ligase (EC 6.2.1.25)
Benzoyl-CoA oxygenase component A
Benzoyl-CoA oxygenase component B
Benzoylformate decarboxylase (EC 4.1.1.7)
Beta N-acetyl-glucosaminidase (EC 3.2.1.52)
Beta-1,4-galactosyltransferase
Beta-carotene ketolase (EC 1.14.-.-)
Beta-fimbriae chaperone protein
Beta-fimbriae probable major subunit
Beta-galactosidase (EC 3.2.1.23)
Beta-glucosidase (EC 3.2.1.21)
Beta-glucoside bgl operon antiterminator, BglG family
Beta-glucuronidase (EC 3.2.1.31)
Beta-hexosaminidase (EC 3.2.1.52)
Beta-ketoadipate enol-lactone hydrolase (EC 3.1.1.24)
Beta-ketoadipyl CoA thiolase (EC 2.3.1.-)
Beta-lactamase
Beta-lactamase (EC 3.5.2.6)
Beta-lactamase class C and other penicillin binding proteins
Beta-lactamase class D
Beta-lactamase regulatory sensor-transducer BlaR1
Beta-lactamase repressor BlaI

Catalase (EC 1.11.1.6)
Catalyzes the cleavage of p-aminobenzoyl-glutamate to p-aminobenzoate and glutamate, subunit A
Catechol 1,2-dioxygenase (EC 1.13.11.1)
Catechol 2,3-dioxygenase (EC 1.13.11.2)
Cation efflux system protein CusC precursor
CblY, a non-orthologous displasment for Alpha-ribazole-5'-phosphate phosphatase
CblZ, a non-orthologous displasment for Alpha-ribazole-5'-phosphate phosphatase
Ccs1/ResB-related putative cytochrome C-type biogenesis protein
CDP-diacylglycerol--glycerol-3-phosphate 3-phosphatidyltransferase (EC 2.7.8.5)
CDP-diacylglycerol--serine O-phosphatidyltransferase (EC 2.7.8.8)
CDP-glycerol: N-acetyl-beta-D-mannosaminyl-1,4-N-acetyl-D-glucosaminyldiphosphoundecaprenyl glycerophosphotransferase
Cell division protein DivIC (FtsB), stabilizes FtsL against RasP cleavage
Cell division protein FtsA
Cell division protein FtsI [Peptidoglycan synthetase] (EC 2.4.1.129)
Cell division protein FtsK
Cell division protein FtsN
Cell division protein FtsQ
Cell division protein FtsQ homolog
Cell division protein FtsW
Cell division protein FtsX
Cell division protein FtsZ (EC 3.4.24.-)
Cell division protein ZipA
Cell division transporter, ATP-binding protein FtsE (TC 3.A.5.1.1)
Cell envelope-associated transcriptional attenuator LytR-CpsA-Psr, subfamily F2 (as in PMID19099556)
CFA/I fimbrial minor adhesin

Heat shock protein GrpE
Heavy metal RND efflux outer membrane protein, CzcC family
Heavy metal sensor histidine kinase
Heme ABC transporter (Streptococcus), ATP-binding protein
Heme ABC transporter (Streptococcus), heme and hemoglobin-binding protein SiaA/HtsA
Heme ABC transporter (Streptococcus), permease protein
Heme ABC type transporter HtsABC, heme-binding protein
Heme ABC type transporter HtsABC, permease protein HtsC
Heme efflux system permease HrtB
Heme oxygenase (EC 1.14.99.3)
Heme transporter IsdDEF, lipoprotein IsdE
Heme transporter IsdDEF, membrane component IsdD
Heme-degrading cytoplasmic oxygenase IsdG
Hemin transport protein HmuS
Hemoglobin-dependent two component system response regulator HrrA
Hemoglobin, heme-dependent two component system sensory histidine kinase ChrS
Hemophore HasA
heterodisulfide reductase, subunit A/methylviologen reducing hydrogenase, subunit delta
Hexokinase (EC 2.7.1.1)
Hexose phosphate uptake regulatory protein UhpC
High-affinity leucine-specific transport system, periplasmic binding protein LivK (TC 3.A.1.4.1)
Histidine ABC transporter, ATP-binding protein HisP (TC 3.A.1.3.1)
Histidine ABC transporter, histidine-binding periplasmic protein precursor HisJ (TC 3.A.1.3.1)
Histidine Biosynthesis in Streptococci
Hmc operon transcriptional regulator Rrf2

Beta-mannosidase (EC 3.2.1.25)
Beta-phosphoglucomutase (EC 5.4.2.6)
Beta-ureidopropionase (EC 3.5.1.6)
Beta-xylosidase (EC 3.2.1.37)
Betaine aldehyde dehydrogenase (EC 1.2.1.8)
Biopolymer transport protein ExbD/TolR
Biosynthetic arginine decarboxylase (EC 4.1.1.19)
Biosynthetic Aromatic amino acid aminotransferase alpha (EC 2.6.1.57)
Biosynthetic Aromatic amino acid aminotransferase beta (EC 2.6.1.57)
Biotin carboxyl carrier protein of acetyl-CoA carboxylase
Biotin carboxylase of acetyl-CoA carboxylase (EC 6.3.4.14)
Biotin synthase (EC 2.8.1.6)
Biphenyl-2,3-diol 1,2-dioxygenase (EC 1.13.11.39)
biphenyl-2,3-diol 1,2-dioxygenase III-related protein
Bis(5'-nucleosyl)-tetraphosphatase (asymmetrical) (EC 3.6.1.17)
Blr3520 protein homolog, hypothetical protein
Blue copper oxidase CueO precursor
Branched-chain acyl-CoA dehydrogenase (EC 1.3.99.12)
Branched-chain alpha-keto acid dehydrogenase, E1 component, alpha subunit (EC 1.2.4.4)
Branched-chain alpha-keto acid dehydrogenase, E1 component, beta subunit (EC 1.2.4.4)
Branched-chain amino acid ABC transporter, amino acid-binding protein (TC 3.A.1.4.1)
Branched-chain amino acid aminotransferase (EC 2.6.1.42)
Branched-chain amino acid transport ATP-binding protein LivF (TC 3.A.1.4.1)
Branched-chain amino acid transport ATP-binding protein LivG (TC 3.A.1.4.1)
Branched-chain amino acid transport system permease protein LivM (TC 3.A.1.4.1)

Chaperone protein DnaJ
Chaperone protein DnaK
Chaperone protein HscA
Chaperone protein HscB
Chaperone protein HtpG
Chemotaxis protein CheV (EC 2.7.3.-)
Chemotaxis protein methyltransferase CheR (EC 2.1.1.80)
Chemotaxis regulator - transmits chemoreceptor signals to flagelllar motor components CheY
Chemotaxis response - phosphatase CheZ
Chemotaxis response regulator protein-glutamate methylesterase CheB (EC 3.1.1.61)
Chitinase (EC 3.2.1.14)
ChlD component of cobalt chelatase involved in B12 biosynthesis
ChlI component of cobalt chelatase involved in B12 biosynthesis
Chloride channel protein
Choline dehydrogenase (EC 1.1.99.1)
Choline-sulfatase (EC 3.1.6.6)
Choline-sulfatase (EC 3.1.6.6)
Chorismate mutase I (EC 5.4.99.5)
Chorismate synthase (EC 4.2.3.5)
Chorismate--pyruvate lyase (EC 4.1.3.40)
Chromate transport protein ChrA
Chromosomal replication initiator protein DnaA
Chromosome (plasmid) partitioning protein ParA
Chromosome (plasmid) partitioning protein ParB
Chromosome partition protein MukB
Chromosome partition protein smc
Circadian clock protein KaiC
Citrate lyase beta chain (EC 4.1.3.6)
Holo-[acyl-carrier protein] synthase (EC 2.7.8.7)
HTH-type transcriptional regulator PtxR
HupE-UreJ family cobalt transporter
Hut operon positive regulatory protein
hydrogenase, methyl-violgen-reducing type, delta subunit
hydrogenase/sulfur reductase, delta subunit
hydrolase, haloacid dehalogenase-like family protein BCZK2594
hydroxycinnamate specific porine
Hydroxymethylglutaryl-CoA lyase (EC 4.1.3.4)
Hypothetical ATP-binding protein, containing DUF265 domain
Hypothetical bd-type Menaquinol oxidase subunit
Hypothetical distantly related to thiol:disulfide interchange protein DsbA
Hypothetical flavoprotein YqcA (clustered with tRNA pseudouridine synthase C)
Hypothetical glycoside hydrolase, family 43, similar to arabinosidase
Hypothetical mga-associated protein
Hypothetical protein (associated with DNA helicase - Rad25 homolog)
Hypothetical protein associated with short form of CcmH
hypothetical protein Bcep3774, commonly clustered with carotenoid biosynthesis
hypothetical protein clusted with conjugative transposons, BF0131
Hypothetical protein distantly related to putative heme lyase CcmC
Hypothetical protein pVir0007
Hypothetical protein pVir0008
Hypothetical protein pVir0009
Hypothetical protein pVir0010
Hypothetical protein pVir0015
Hypothetical protein pVir0016
Hypothetical protein pVir0019/pVir0020
Hypothetical protein pVir0021
Butyryl-CoA dehydrogenase (EC 1.3.99.2)
C-terminal domain of CinA type E
C-terminal domain of CinA type S
CAMP phosphodiesterases class-II:Metallo-beta-lactamase superfamily
cAMP-binding proteins - catabolite gene activator and regulatory subunit of cAMP-dependent protein kinases
cAMP-dependent Kef-type K+ transport system
Capsular polysaccharide biosynthesis protein CapD
Carbamate kinase (EC 2.7.2.2)
Carbamoyl-phosphate synthase large chain (EC 6.3.5.5)
Carbamoyl-phosphate synthase small chain (EC 6.3.5.5)
carbon monoxide dehydrogenase D protein
carbon monoxide dehydrogenase E protein
Carbon monoxide dehydrogenase F protein
Carbon monoxide dehydrogenase large chain (EC 1.2.99.2)
Carbon monoxide dehydrogenase medium chain (EC 1.2.99.2)
Carbon monoxide dehydrogenase small chain (EC 1.2.99.2)
Carbon starvation protein A
Carbon-monoxide dehydrogenase form II, large subunit (EC 1.2.99.2)
Carbonic anhydrase (EC 4.2.1.1)
Cardiolipin synthetase (EC 2.7.8.-)
Carnitine racemase (EC 5.-.-.-)
Catalase (EC 1.11.1.6)
Catalyzes the cleavage of p-aminobenzoyl-glutamate to p-aminobenzoate and glutamate, subunit A
Catechol 1,2-dioxygenase (EC 1.13.11.1)
Catechol 1,2-dioxygenase 1 (EC 1.13.11.1)
Cation efflux system protein CusC precursor
CblZ, a non-orthologous displasment for Alpha-ribazole-5'-phosphate phosphatase
Ccs1/ResB-related putative cytochrome C-type biogenesis protein

Citrate synthase (si) (EC 2.3.3.1)

ClpB protein

CoA-acylating propionaldehyde dehydrogenase

Cob(I)alamin adenosyltransferase (EC 2.5.1.17)

Cob(I)alamin adenosyltransferase PduO (EC 2.5.1.17)

Cobalamin synthase

Cobalt-precorrin-2 C20-methyltransferase (EC 2.1.1.130)

Cobalt-precorrin-3b C17-methyltransferase

Cobalt-precorrin-4 C11-methyltransferase (EC 2.1.1.133)

Cobalt-precorrin-6x reductase (EC 1.3.1.54)

Cobalt-precorrin-6y C5-methyltransferase (EC 2.1.1.-)

Cobalt-precorrin-8x methylmutase (EC 5.4.1.2)

Cobalt-zinc-cadmium resistance protein

Cobalt-zinc-cadmium resistance protein CzcA

Cobalt-zinc-cadmium resistance protein CzcD

Cobalt/zinc/cadmium efflux RND transporter, membrane fusion protein, CzcB family

CobN component of cobalt chelatase involved in B12 biosynthesis

Cobyrinic acid A,C-diamide synthase

Coenzyme PQQ synthesis protein B

Coenzyme PQQ synthesis protein D

Coenzyme PQQ synthesis protein E

COG0488: ATPase components of ABC transporters with duplicated ATPase domains

COG0613, Predicted metal-dependent phosphoesterases (PHP family)

COG4123: Predicted O-methyltransferase

Colanic acid biosynthesis glycosyl transferase WcaA

Colanic acid polymerase WcaD

Competence protein F homolog, phosphoribosyltransferase domain

Hypothetical protein pVir0029, glutamine-rich

Hypothetical protein pVir0042

Hypothetical protein, Lin0079 homolog [Bacteriophage A118]

Hypothetical protein, Lmo2307 homolog [Bacteriophage A118]

Hypothetical similar to CcmC, putative heme lyase for CcmE

Hypothetical succinate dehydrogenase membrane anhor protein

Hypothetical, distant similarity with heme-degrading oxygenase IsdG

IncF plasmid conjugative transfer mating signal transduction protein TraM

IncF plasmid conjugative transfer pilin acetylase TraX

IncF plasmid conjugative transfer pilin protein TraA

IncF plasmid conjugative transfer pilus assembly protein TraB

IncF plasmid conjugative transfer pilus assembly protein TraC

IncF plasmid conjugative transfer pilus assembly protein TraE

IncF plasmid conjugative transfer pilus assembly protein TraH

IncF plasmid conjugative transfer pilus assembly protein TraK

IncF plasmid conjugative transfer pilus assembly protein TraL

IncF plasmid conjugative transfer pilus assembly protein TraU

IncF plasmid conjugative transfer pilus assembly protein TraV

IncF plasmid conjugative transfer pilus assembly protein TraW

IncF plasmid conjugative transfer protein TraD

IncF plasmid conjugative transfer protein TraG

IncF plasmid conjugative transfer protein TraN

IncF plasmid conjugative transfer protein TraQ

IncF plasmid conjugative transfer protein TraR

IncF plasmid conjugative transfer protein TrbA

IncF plasmid conjugative transfer protein TrbC

IncF plasmid conjugative transfer protein TrbD

CDP-diacylglycerol--glycerol-3-phosphate 3-phosphatidyltransferase (EC 2.7.8.5)

CDP-diacylglycerol--serine O-phosphatidyltransferase (EC 2.7.8.8)

Cell division protein FtsA

Cell division protein FtsI [Peptidoglycan synthetase] (EC 2.4.1.129)

Cell division protein FtsK

Cell division protein FtsL

Cell division protein FtsQ

Cell division protein FtsW

Cell division protein FtsX

Cell division protein FtsZ (EC 3.4.24.-)

Cell division protein MraZ

Cell division transporter, ATP-binding protein FtsE (TC 3.A.5.1.1)

Cell envelope-associated transcriptional attenuator LytR-CpsA-Psr, subfamily A1 (as in PMID19099556)

Cell envelope-associated transcriptional attenuator LytR-CpsA-Psr, subfamily M (as in PMID19099556)

Cellobiose phosphotransferase system YdjC-like protein

CFA/I fimbrial minor adhesin

Chaperone protein DnaJ

Chaperone protein DnaK

Chaperone protein HscA

Chaperone protein HscB

Chaperone protein HtpG

Chemotaxis protein CheV (EC 2.7.3.-)

Chemotaxis protein methyltransferase CheR (EC 2.1.1.80)

Chemotaxis regulator - transmits chemoreceptor signals to flagelllar motor components CheY

Chemotaxis response - phosphatase CheZ

Chemotaxis response regulator protein-glutamate methylesterase CheB (EC 3.1.1.61)

Chitinase (EC 3.2.1.14)

Conjugative transfer protein TrbB	IncF plasmid conjugative transfer protein TrbE	ChlD component of cobalt chelatase involved in B12 biosynthesis
Conjugative transfer protein TrbC	IncF plasmid conjugative transfer protein TrbG	ChlI component of cobalt chelatase involved in B12 biosynthesis
Conjugative transfer protein TrbE	IncF plasmid conjugative transfer protein TrbH	Chloride channel protein
Conjugative transfer protein TrbF	IncF plasmid conjugative transfer protein TrbI	Choline dehydrogenase (EC 1.1.99.1)
Conjugative transfer protein TrbI	IncF plasmid conjugative transfer protein TrbJ	Choline-sulfatase (EC 3.1.6.6)
Conjugative transfer protein TrbJ	IncF plasmid conjugative transfer regulator TraJ	Chorismate mutase I (EC 5.4.99.5)
Conjugative transfer protein TrbL	IncF plasmid conjugative transfer surface exclusion protein TraS	Chorismate synthase (EC 4.2.3.5)
Conjugative transposon protein TraE	IncF plasmid conjugative transfer surface exclusion protein TraT	Chorismate--pyruvate lyase (EC 4.1.3.40)
Conjugative transposon protein TraG	IncW plasmid conjugative relaxase protein TrwC (TraI homolog)	Chromate transport protein ChrA
Conjugative transposon protein TraI	Indolepyruvate oxidoreductase subunit IorB (EC 1.2.7.8)	Chromosomal replication initiator protein DnaA
Conjugative transposon protein TraJ	Inhibitor of vertebrate lysozyme precursor	Chromosome (plasmid) partitioning protein ParA
Conjugative transposon protein TraK	Inner membrane protein forms channel for type IV secretion of T-DNA complex (VirB8)	Chromosome (plasmid) partitioning protein ParB
Conjugative transposon protein TraM	Inositol-1-phosphate synthase (EC 5.5.1.4)	Chromosome partition protein MukB
Conjugative transposon protein TraO	Ins1, sterol regulated ER retention factor for SREBP-SCAP complexes	Chromosome partition protein MukE
Conserved hypothetical protein (perhaps related to histidine degradation)	Insig-2, sterol regulated ER retention factor for SREBP-SCAP complexes	Chromosome partition protein smc
Conserved protein YghR, with nucleoside triphosphate hydrolase domain	Internalin C	CidA-associated membrane protein CidB
Conserved uncharacterized protein CreA	Internalin H (LPXTG motif)	Citrate lyase beta chain (EC 4.1.3.6)
CopG protein	Internalin-like protein (LPXTG motif) Lin0372 homolog	Citrate synthase (si) (EC 2.3.3.1)
Copper resistance protein B	Internalin-like protein (LPXTG motif) Lin1204 homolog	ClpB protein
Copper sensory histidine kinase CpxA	Internalin-like protein (LPXTG motif) Lmo0327 homolog	Co-activator of prophage gene expression IbrA
Copper sensory histidine kinase CusS	Internalin-like protein (LPXTG motif) Lmo0331 homolog	Co-activator of prophage gene expression IbrB
Copper tolerance protein	Internalin-like protein (LPXTG motif) Lmo0409 homolog	CoA-acylating propionaldehyde dehydrogenase
Copper-containing nitrite reductase (EC 1.7.2.1)	Internalin-like protein (LPXTG motif) Lmo1290 homolog	Cob(I)alamin adenosyltransferase (EC 2.5.1.17)
Copper-sensing two-component system response regulator CusR	Internalin-like protein (LPXTG motif) Lmo2396 homolog	Cob(I)alamin adenosyltransferase PduO (EC 2.5.1.17)
Copper-translocating P-type ATPase (EC 3.6.3.4)	Internalin-like protein (LPXTG motif) Lmo2821 homolog	Cobalamin biosynthesis protein BluB
Coproporphyrinogen III oxidase, oxygen-independent (EC 1.3.99.22)	internalin, putative	Cobalamin synthase
Creatinine amidohydrolase (EC 3.5.2.10)	Iron compound ABC uptake transporter ATP-binding protein	Cobalt-precorrin-2 C20-methyltransferase (EC 2.1.1.130)
CRISPR-associated helicase Cas3	Iron compound ABC uptake transporter ATP-binding protein PiaD	Cobalt-precorrin-3b C17-methyltransferase
CRISPR-associated helicase Cas3, protein	Iron compound ABC uptake transporter permease protein PiaB	Cobalt-precorrin-4 C11-methyltransferase (EC 2.1.1.133)

CRISPR-associated protein, Cse3 family
Crossover junction endodeoxyribonuclease RuvC (EC 3.1.22.4)
Crotonobetainyl-CoA:carnitine CoA-transferase (EC 2.8.3.-)
CsdL (EC-YgdL) protein of the HesA/MoeB/ThiF family, part of the CsdA-E-L sulfur transfer pathway
CTP synthase (EC 6.3.4.2)
CTP:molybdopterin cytidylyltransferase
Cyanate ABC transporter, substrate binding protein
Cyanate hydratase (EC 4.2.1.104)
Cyclic AMP receptor protein
Cyclic pyranopterin phosphate synthase (MoaA) (EC 4.1.99.18)
Cyclohexadienyl dehydratase (EC 4.2.1.51)(EC 4.2.1.91)
Cyclohexadienyl dehydrogenase (EC 1.3.1.12)(EC 1.3.1.43)
Cyclohexanecarboxylate-CoA ligase AliA (EC 6.2.1.-)
Cys regulon transcriptional activator CysB
Cystathionine beta-lyase (EC 4.4.1.8)
Cystathionine beta-synthase (EC 4.2.1.22)
Cystathionine gamma-lyase (EC 4.4.1.1)
Cystathionine gamma-synthase (EC 2.5.1.48)
Cysteine desulfurase (EC 2.8.1.7)
Cysteine desulfurase (EC 2.8.1.7), IscS subfamily
Cysteine desulfurase (EC 2.8.1.7), NifS subfamily
Cysteine desulfurase (EC 2.8.1.7), SufS subfamily
Cysteine desulfurase CsdA-CsdE (EC 2.8.1.7), main protein CsdA
Cysteine synthase (EC 2.5.1.47)
Cysteine synthase B (EC 2.5.1.47)
Cysteinyl-tRNA synthetase (EC 6.1.1.16)
Cytidine deaminase (EC 3.5.4.5)
Cytochrome c heme lyase subunit CcmF
Cytochrome c oxidase (B(O/a)3-type) chain I (EC 1.9.3.1)
Cytochrome c oxidase (B(O/a)3-type) chain II (EC 1.9.3.1)

Iron compound ABC uptake transporter permease protein PiaC
Iron compound ABC uptake transporter permease protein PiuC
Iron compound ABC uptake transporter substrate-binding protein
Iron compound ABC uptake transporter substrate-binding protein PiaA
iron-chelator utilization protein
Iron-dependent repressor IdeR/DtxR
Iron-responsive repressor RirA
Iron-sulfur cluster assembly scaffold protein IscU
Iron-sulfur cluster regulator IscR
Isochorismatase (EC 3.3.2.1) of siderophore biosynthesis
Isopentenyl phosphate kinase
Isopentenyl-diphosphate delta-isomerase, FMN-dependent (EC 5.3.3.2)
Ketoisovalerate oxidoreductase subunit VorG (EC 1.2.7.7)
L-allo-threonine aldolase (EC 2.1.2.1)
L-arabinolactonase (EC 3.1.1.15)
L-aspartate oxidase (EC 1.4.3.16)
L-Proline/Glycine betaine transporter ProP
L-rhamnonate dehydratase (EC 4.2.1.90)
L-rhamnose isomerase (EC 5.3.1.14)
L-rhamnose operon transcriptional activator RhaR
L-rhamnose-1-dehydrogenase (EC 1.1.1.173)
L-ribulose-5-phosphate 4-epimerase UlaF (EC 5.1.3.4) (L-ascorbate utilization protein F)
L-sorbose 1-phosphate reductase (EC 1.1.1.-)
L-threonine transporter, anaerobically inducible
L-xylulose 5-phosphate 3-epimerase (EC 5.1.3.-)
Lactate 2-monooxygenase (EC 1.13.12.4)
Lactose and galactose permease, GPH translocator family
Lactose permease
Large-conductance mechanosensitive channel
Lipoate synthase

Cobalt-precorrin-6x reductase (EC 1.3.1.54)
Cobalt-precorrin-6y C15-methyltransferase [decarboxylating] (EC 2.1.1.-)
Cobalt-precorrin-6y C5-methyltransferase (EC 2.1.1.-)
Cobalt-zinc-cadmium resistance protein
Cobalt-zinc-cadmium resistance protein CzcA
Cobalt-zinc-cadmium resistance protein CzcD
Cobalt/zinc/cadmium efflux RND transporter, membrane fusion protein, CzcB family
CobN component of cobalt chelatase involved in B12 biosynthesis
Coenzyme PQQ synthesis protein B
Coenzyme PQQ synthesis protein D
Coenzyme PQQ synthesis protein E
COG0488: ATPase components of ABC transporters with duplicated ATPase domains
COG0613, Predicted metal-dependent phosphoesterases (PHP family)
COG3178: Predicted phosphotransferase related to Ser/Thr protein kinases
Colanic acid biosynthesis glycosyl transferase WcaC
Colanic acid biosynthesis glycosyl transferase WcaL
Colanic acid biosynthsis UDP-glucose lipid carrier transferase WcaJ
Colanic acid biosysnthesis protein WcaK
Colanic acid polymerase WcaD
Cold shock protein CspE
Colicin I receptor precursor
Competence protein F homolog, phosphoribosyltransferase domain
Conjugative signal peptidase TrhF
Conjugative transfer protein TrbA
Conjugative transfer protein TrbB
Conjugative transfer protein TrbC
Conjugative transfer protein TrbE
Conjugative transfer protein TrbF
Conjugative transfer protein TrbG
Conjugative transfer protein TrbI

Cytochrome c oxidase polypeptide I (EC 1.9.3.1)
Cytochrome c oxidase polypeptide II (EC 1.9.3.1)
Cytochrome c oxidase polypeptide III (EC 1.9.3.1)
Cytochrome c oxidase subunit CcoN (EC 1.9.3.1)
Cytochrome c oxidase subunit CcoO (EC 1.9.3.1)
Cytochrome c oxidase subunit CcoP (EC 1.9.3.1)
Cytochrome c oxidase subunit CcoQ (EC 1.9.3.1)
Cytochrome c-type biogenesis protein CcdA (DsbD analog)
Cytochrome c-type biogenesis protein CcmC, putative heme lyase for CcmE
Cytochrome c-type biogenesis protein CcmG/DsbE, thiol:disulfide oxidoreductase
Cytochrome c-type biogenesis protein Ccs1/ResB
Cytochrome c-type biogenesis protein CcsA/ResC
Cytochrome c-type biogenesis protein DsbD, protein-disulfide reductase (EC 1.8.1.8)
Cytochrome c-type protein NapC
Cytochrome c-type protein TorC
Cytochrome c-type protein TorY
Cytochrome c4
Cytochrome C550 (Soluble cytochrome C)
Cytochrome c551 peroxidase (EC 1.11.1.5)
Cytochrome c552 precursor (EC 1.7.2.2)
Cytochrome c553
Cytochrome c553-like
Cytochrome d ubiquinol oxidase subunit I (EC 1.10.3.-)
Cytochrome d ubiquinol oxidase subunit II (EC 1.10.3.-)
Cytochrome O ubiquinol oxidase subunit I (EC 1.10.3.-)
Cytochrome O ubiquinol oxidase subunit II (EC 1.10.3.-)
Cytochrome O ubiquinol oxidase subunit III (EC 1.10.3.-)
Cytochrome oxidase biogenesis protein Sco1/SenC/PrrC, putative copper metallochaperone
Cytochrome oxidase biogenesis protein Surf1, facilitates heme A insertion
Cytoplasmic axial filament protein CafA and Ribonuclease G (EC 3.1.4.-)

Lipoate synthase, cyanobacterial paralog
Lipoate-protein ligase A
Lipoate-protein ligase A, N-terminal 30 percent
Lipopolysaccharide cholinephosphotransferase LicD2 (EC 2.7.8.-)
Low molecular weight protein tyrosine phosphatase (EC 3.1.3.48)
LPXTG specific sortase A
LsrR, transcriptional repressor of lsr operon
Lumazine protein, riboflavin synthase homolog
LuxG, NAD(P)H-dependent FMN reductase (EC 1.5.1.29)
M protein trans-acting positive regulator (Mga)
Macrolide-specific efflux protein MacA
Major capsid protein
Major curlin subunit precursor CsgA
Malate synthase G (EC 2.3.3.9)
Malolactic regulator
Malonate decarboxylase alpha subunit
Malonate decarboxylase beta subunit
Malonate decarboxylase delta subunit
Malonate decarboxylase gamma subunit
Malonate transporter, MadL subunit
Malonate transporter, MadM subunit
Malonate utilization transcriptional regulator
Malonyl CoA-acyl carrier protein transacylase (EC 2.3.1.39)
Malto-oligosyltrehalose trehalohydrolase (EC 3.2.1.141)
Maltose/maltodextrin ABC transporter, substrate binding periplasmic protein MalE
Manganese superoxide dismutase (EC 1.15.1.1)
Mannose-1-phosphate guanylyltransferase (EC 2.7.7.13)
Mannose-1-phosphate guanylyltransferase (GDP) (EC 2.7.7.22)
Mannose-6-phosphate isomerase, class I (EC 5.3.1.8)
Mannosides-regulated TonB-dependent outer membrane receptor

Conjugative transfer protein TrbJ
Conjugative transfer protein TrbL
Conjugative transposon protein TraM
Conserved protein YcjX with nucleoside triphosphate hydrolase domain
Conserved protein YghR, with nucleoside triphosphate hydrolase domain
Conserved protein YghT, with nucleoside triphosphate hydrolase domain
Conserved uncharacterized protein CreA
CopG protein
Copper homeostasis protein CutF precursor
Copper resistance protein B
Copper sensory histidine kinase CpxA
Copper sensory histidine kinase CpxA
Copper sensory histidine kinase CpxA
Copper sensory histidine kinase CusS
Copper sensory histidine kinase CusS
Copper sensory histidine kinase CusS
Copper tolerance protein
Copper-containing nitrite reductase (EC 1.7.2.1)
Copper-sensing two-component system response regulator CusR
Copper-translocating P-type ATPase (EC 3.6.3.4)
Coproporphyrinogen III oxidase, oxygen-independent (EC 1.3.99.22)
CRISPR-associated helicase Cas3
CRISPR-associated helicase Cas3, protein
CRISPR-associated protein Cas1
CRISPR-associated protein TM1812
CRISPR-associated protein, Cas5e family
CRISPR-associated protein, Csd1 family
CRISPR-associated protein, Csd2/Csh2 family
CRISPR-associated protein, Cse3 family
CRISPR-associated protein, Cse4 family

Cytoplasmic trehalase (EC 3.2.1.28)
Cytosine/purine/uracil/thiamine/allantoin permease family protein
cytosolic long-chain acyl-CoA thioester hydrolase family protein
D-3-phosphoglycerate dehydrogenase (EC 1.1.1.95)
D-alanine aminotransferase (EC 2.6.1.21)
D-alanine--D-alanine ligase (EC 6.3.2.4)
D-alanine--D-alanine ligase A (EC 6.3.2.4)
D-alanyl transfer protein DltB
D-alanyl-D-alanine carboxypeptidase (EC 3.4.16.4)
D-alanyl-D-alanine carboxypeptidase (EC 3.4.16.4)
D-allose ABC transporter, substrate-binding component
D-amino acid dehydrogenase small subunit (EC 1.4.99.1)
D-arabino-3-hexulose 6-phosphate formaldehyde-lyase (EC 4.1.2.43)
D-beta-hydroxybutyrate dehydrogenase (EC 1.1.1.30)
D-beta-hydroxybutyrate permease
D-erythrose-4-phosphate dehydrogenase (EC 1.2.1.72)
D-galactarate dehydratase (EC 4.2.1.42)
D-galactarate permease
D-galactonate transporter
D-glucarate permease
D-glycerate 2-kinase (EC 2.7.1.-)
D-glycerate transporter (predicted)
D-glycero-beta-D-manno-heptose 7-phosphate kinase
D-hydantoinase (EC 3.5.2.2)
D-lactate dehydrogenase (EC 1.1.1.28)
D-Lactate dehydrogenase, cytochrome c-dependent (EC 1.1.2.4)
D-malic enzyme (EC 1.1.1.83)
D-mannonate oxidoreductase (EC 1.1.1.57)
Membrane protein, suppressor for copper-sensitivity ScsB
Menaquinone via futalosine step 2
Menaquinone-specific isochorismate synthase (EC 5.4.4.2)
Metal-dependent hydrolase involved in phosphonate metabolism
Metallo-beta-lactamase superfamily domain protein in prophage
Metalloprotease, putative zinc-binding domain
Methanophenazine hydrogenase maturation protease (EC 3.4.24.-)
Methyl coenzyme M reductase gamma subunit (EC 2.8.4.1)
Methyl coenzyme M reductase I alpha subunit (EC 2.8.4.1)
Methyl coenzyme M reductase I beta subunit (EC 2.8.4.1)
Methyl coenzyme M reductase I gamma subunit (EC 2.8.4.1)
Methyl coenzyme M reductase II alpha subunit (EC 2.8.4.1)
Methyl coenzyme M reductase II gamma subunit (EC 2.8.4.1)
methyl viologen-reducing hydrogenase, delta subunit homolog FlpD
Methyl-accepting chemotaxis protein I (serine chemoreceptor protein)
Methyl-accepting chemotaxis protein III (ribose and galactose chemoreceptor protein)
Methyl-accepting chemotaxis protein IV (dipeptide chemoreceptor protein)
Methylcrotonyl-CoA carboxylase biotin-containing subunit (EC 6.4.1.4)
Methylmalonyl-CoA decarboxylase, beta chain (EC 4.1.1.41)
Methylthioribose-1-phosphate isomerase (EC 5.3.1.23)
Mga-associated protein
Minor curlin subunit CsgB, nucleation component of curlin monomers
Molybdate-binding domain of ModE
Molybdenum ABC transporter, periplasmic molybdenum-binding protein ModA (TC 3.A.1.8.1)
Molybdenum cofactor biosynthesis protein MoaB
Molybdenum cofactor biosynthesis protein MoaC
Molybdenum cofactor biosynthesis protein MoaD
Molybdenum cofactor biosynthesis protein MoaE
CRISPR-associated protein, Csn1 family
CRISPR-associated RAMP Cmr4
Crossover junction endodeoxyribonuclease RuvC (EC 3.1.22.4)
Crotonobetainyl-CoA:carnitine CoA-transferase (EC 2.8.3.-)
CTP synthase (EC 6.3.4.2)
CTP:molybdopterin cytidylyltransferase
Curli production assembly/transport component CsgF
Cyanate ABC transporter, ATP-binding protein
Cyanate ABC transporter, permease protein
Cyanate ABC transporter, substrate binding protein
Cyanate transport protein CynX
Cyclic pyranopterin phosphate synthase (MoaA) (EC 4.1.99.18)
Cyclohexadienyl dehydratase (EC 4.2.1.51)(EC 4.2.1.91)
Cyclohexadienyl dehydrogenase (EC 1.3.1.12)(EC 1.3.1.43)
Cyn operon transcriptional activator
Cys regulon transcriptional activator CysB
Cystathionine beta-lyase (EC 4.4.1.8)
Cystathionine gamma-lyase (EC 4.4.1.1)
Cystathionine gamma-synthase (EC 2.5.1.48)
Cysteine desulfurase (EC 2.8.1.7)
Cysteine desulfurase (EC 2.8.1.7), IscS subfamily
Cysteine desulfurase (EC 2.8.1.7), NifS subfamily
Cysteine desulfurase (EC 2.8.1.7), SufS subfamily
Cysteine desulfurase CsdA-CsdE (EC 2.8.1.7), main protein CsdA
Cysteine desulfurase CsdA-CsdE, sulfur acceptor protein CsdE
Cysteine synthase (EC 2.5.1.47)
Cysteine synthase B (EC 2.5.1.47)
Cysteinyl-tRNA synthetase (EC 6.1.1.16)

D-mannose isomerase (EC 5.3.1.7)	Molybdenum transport ATP-binding protein ModC (TC 3.A.1.8.1)	Cystine ABC transporter, periplasmic cystine-binding protein FliY
D-serine dehydratase transcriptional activator	Molybdenum transport system permease protein ModB (TC 3.A.1.8.1)	Cytidine deaminase (EC 3.5.4.5)
D-serine/D-alanine/glycine transporter	Molybdenum transport system protein ModD	Cytidylate kinase (EC 2.7.4.25)
D-xylose 1-dehydrogenase (EC 1.1.1.175)	Molybdopterin biosynthesis Mog protein, molybdochelatase	Cytochrome c heme lyase subunit CcmF
D-xylose proton-symporter XylE	Molybdopterin biosynthesis protein MoeA	Cytochrome c heme lyase subunit CcmH
D-xylose transport ATP-binding protein XylG	Molybdopterin biosynthesis protein MoeB	Cytochrome c heme lyase subunit CcmL
D,D-heptose 7-phosphate kinase	Molybdopterin-guanine dinucleotide biosynthesis protein MobA	Cytochrome c oxidase (B(O/a)3-type) chain I (EC 1.9.3.1)
DEAD-box ATP-dependent RNA helicase CshA (EC 3.6.4.13)	Molybdopterin-guanine dinucleotide biosynthesis protein MobB	Cytochrome c oxidase polypeptide I (EC 1.9.3.1)
Death on curing protein, Doc toxin	Monofunctional biosynthetic peptidoglycan transglycosylase (EC 2.4.2.-)	Cytochrome c oxidase polypeptide II (EC 1.9.3.1)
Decaprenyl diphosphate synthase (EC 2.5.1.91)	Monoglyceride lipase (EC 3.1.1.23)	Cytochrome c oxidase polypeptide III (EC 1.9.3.1)
Delta-1-pyrroline-5-carboxylate dehydrogenase (EC 1.5.1.12)	Monomethylamine methyltransferase corrinoid protein	Cytochrome c oxidase subunit CcoN (EC 1.9.3.1)
Deoxyadenosine kinase (EC 2.7.1.76)	Monomethylamine permease	Cytochrome c oxidase subunit CcoO (EC 1.9.3.1)
Deoxycytidine triphosphate deaminase (EC 3.5.4.13)	Monomethylamine:corrinoid methyltransferase	Cytochrome c oxidase subunit CcoP (EC 1.9.3.1)
Deoxyguanosinetriphosphate triphosphohydrolase (EC 3.1.5.1)	MSHA biogenesis protein MshE	Cytochrome c-552 precursor
Deoxyribodipyrimidine photolyase (EC 4.1.99.3)	MSHA biogenesis protein MshK2	Cytochrome c-type biogenesis protein CcdA (DsbD analog)
Deoxyribonucleoside regulator DeoR (transcriptional repressor)	MSHA biogenesis protein MshK3	Cytochrome c-type biogenesis protein CcmC, putative heme lyase for CcmE
Deoxyribose operon repressor, DeoR family	MSHA biogenesis protein MshL	Cytochrome c-type biogenesis protein CcmD, interacts with CcmCE
Deoxyribose-phosphate aldolase (EC 4.1.2.4)	MSHA biogenesis protein MshN	Cytochrome c-type biogenesis protein CcmE, heme chaperone
Deoxyuridine 5'-triphosphate nucleotidohydrolase (EC 3.6.1.23)	MSHA biogenesis protein MshO	Cytochrome c-type biogenesis protein CcmG/DsbE, thiol:disulfide oxidoreductase
Dephospho-CoA kinase (EC 2.7.1.24)	MSHA pilin protein MshD	Cytochrome c-type biogenesis protein Ccs1/ResB
Dethiobiotin synthetase (EC 6.3.3.3)	Multidrug efflux RND membrane fusion protein MexC	Cytochrome c-type biogenesis protein CcsA/ResC
Di/tripeptide permease YjdL	Multimodular transpeptidase-transglycosylase (EC 2.4.1.129) (EC 3.4.-.-)	Cytochrome c-type biogenesis protein DsbD, protein-disulfide reductase (EC 1.8.1.8)
Diacylglycerol kinase (EC 2.7.1.107)	Multiple polyol-specific dehydrogenase (EC 1.1.1.-)	Cytochrome c-type biogenesis protein ResA
Diaminohydroxyphosphoribosylaminopyrimidine deaminase (EC 3.5.4.26)	Multiple virulence factor regulator MvfR/PqsR	Cytochrome c-type heme lyase subunit nrfE, nitrite reductase complex assembly
Diaminopimelate decarboxylase (EC 4.1.1.20)	Myo-inositol 2-dehydrogenase (EC 1.1.1.18)	Cytochrome c-type heme lyase subunit nrfG, nitrite reductase complex assembly
Diaminopimelate epimerase (EC 5.1.1.7)	N-acetylglucosamine-1-phosphate uridyltransferase eukaryotic (EC 2.7.7.23)	Cytochrome c-type protein NapC
diglucosyldiacylglycerol synthase (LTA membrane anchor synthesis)	N-acetylglucosamine-regulated TonB-dependent outer membrane receptor	Cytochrome c-type protein TorC
diguanylate cyclase/phosphodiesterase (GGDEF & EAL domains) with PAS/PAC sensor(s)	N-Acetyltransferase PseH involved in the biosynthesis of pseudaminic acid	Cytochrome c2
Dihydrofolate reductase (EC 1.5.1.3)	N-acyl-L-homoserine lactone synthetase LasI	Cytochrome c4

Dihydrofolate synthase (EC 6.3.2.12)
Dihydrolipoamide acetyltransferase component (E2) of acetoin dehydrogenase complex (EC 2.3.1.-)
Dihydrolipoamide acetyltransferase component of pyruvate dehydrogenase complex (EC 2.3.1.12)
Dihydrolipoamide acyltransferase component of branched-chain alpha-keto acid dehydrogenase complex (EC 2.3.1.168)
Dihydrolipoamide dehydrogenase (EC 1.8.1.4)
Dihydrolipoamide dehydrogenase of 2-oxoglutarate dehydrogenase (EC 1.8.1.4)
Dihydrolipoamide dehydrogenase of branched-chain alpha-keto acid dehydrogenase (EC 1.8.1.4)
Dihydrolipoamide dehydrogenase of pyruvate dehydrogenase complex (EC 1.8.1.4)
Dihydrolipoamide succinyltransferase component (E2) of 2-oxoglutarate dehydrogenase complex (EC 2.3.1.61)
Dihydroneopterin aldolase (EC 4.1.2.25)
Dihydroneopterin triphosphate epimerase
Dihydroorotase (EC 3.5.2.3)
Dihydroorotate dehydrogenase (EC 1.3.3.1)
Dihydroorotate dehydrogenase, catalytic subunit (EC 1.3.3.1)
Dihydropteroate synthase (EC 2.5.1.15)
Dihydropyrimidinase (EC 3.5.2.2)
Dihydropyrimidine dehydrogenase [NADP+] (EC 1.3.1.2)
Dihydroxy-acid dehydratase (EC 4.2.1.9)
Dihydroxyacetone kinase, ATP-dependent (EC 2.7.1.29)
Dimethylhistidine N-methyltransferase (EC 2.1.1.44)
DinG family ATP-dependent helicase CPE1197
DinG family ATP-dependent helicase YoaA
Dipeptide transport ATP-binding protein DppD (TC 3.A.1.5.2)
Dipeptide transport ATP-binding protein DppF (TC 3.A.1.5.2)
Dipeptide transport system permease protein DppB (TC 3.A.1.5.2)

N-acyl-L-homoserine lactone synthetase RhlL
N-formylglutamate deformylase (EC 3.5.1.68)
N-succinyl-L,L-diaminopimelate aminotransferase (EC 2.6.1.17)
N5-methyltetrahydromethanopterin:coenzyme M methyltransferase subunit G (EC 2.1.1.86)
N5-methyltetrahydromethanopterin:coenzyme M methyltransferase subunit H (EC 2.1.1.86)
N6-hydroxylysine O-acetyltransferase (EC 2.3.1.102), aerobactin biosynthesis protein IucB
Na(+)-translocating NADH-quinone reductase subunit D (EC 1.6.5.-)
Na(+)-translocating NADH-quinone reductase subunit F (EC 1.6.5.-)
Na+/H+ antiporter NapA
NAD-reducing hydrogenase subunit HoxY (EC 1.12.1.2)
NADH dehydrogenase (EC 1.6.99.3)
NADH dehydrogenase subunit 1
NADH dehydrogenase subunit 2
NADH dehydrogenase subunit 5
NADH oxidoreductase hcr (EC 1.-.-.-)
NADH-ubiquinone oxidoreductase chain F (EC 1.6.5.3)
NADH-ubiquinone oxidoreductase chain I (EC 1.6.5.3)
NADH-ubiquinone oxidoreductase chain M (EC 1.6.5.3)
Naphthoate synthase (EC 4.1.3.36)
Ni,Fe-hydrogenase I large subunit
Ni,Fe-hydrogenase maturation factor
Nickel responsive regulator NikR
Nicotinamidase (EC 3.5.1.19)
Nicotinate phosphoribosyltransferase (EC 2.4.2.11)
Nitric oxide reductase activation protein NorE

Cytochrome c551 peroxidase (EC 1.11.1.5)
Cytochrome c552 precursor (EC 1.7.2.2)
Cytochrome c553-like
Cytochrome d ubiquinol oxidase subunit I (EC 1.10.3.-)
Cytochrome d ubiquinol oxidase subunit II (EC 1.10.3.-)
Cytochrome O ubiquinol oxidase subunit I (EC 1.10.3.-)
Cytochrome O ubiquinol oxidase subunit II (EC 1.10.3.-)
Cytochrome O ubiquinol oxidase subunit III (EC 1.10.3.-)
Cytochrome O ubiquinol oxidase subunit IV (EC 1.10.3.-)
Cytochrome oxidase biogenesis protein Cox11-CtaG, copper delivery to Cox1
Cytochrome oxidase biogenesis protein Sco1/SenC/PrrC, putative copper metallochaperone
Cytochrome oxidase biogenesis protein Surf1, facilitates heme A insertion
Cytoplasmic alpha-amylase (EC 3.2.1.1)
Cytoplasmic axial filament protein CafA and Ribonuclease G (EC 3.1.4.-)
Cytoplasmic copper homeostasis protein cutC
Cytoplasmic trehalase (EC 3.2.1.28)
Cytosine/purine/uracil/thiamine/allantoin permease family protein
D-3-phosphoglycerate dehydrogenase (EC 1.1.1.95)
D-alanine aminotransferase (EC 2.6.1.21)
D-alanine--D-alanine ligase (EC 6.3.2.4)
D-alanine--D-alanine ligase A (EC 6.3.2.4)
D-alanyl-D-alanine carboxypeptidase (EC 3.4.16.4)
D-allose ABC transporter, ATPase component
D-allose ABC transporter, substrate-binding component
D-amino acid dehydrogenase small subunit (EC 1.4.99.1)

Dipeptide transport system permease protein DppC (TC 3.A.1.5.2)
Dipeptide-binding ABC transporter, periplasmic substrate-binding component (TC 3.A.1.5.2)
Dipeptidyl carboxypeptidase Dcp (EC 3.4.15.5)
DNA gyrase subunit A (EC 5.99.1.3)
DNA gyrase subunit B (EC 5.99.1.3)
DNA helicase IV
DNA ligase (EC 6.5.1.2)
DNA mismatch repair endonuclease MutH
DNA mismatch repair protein MutL
DNA mismatch repair protein MutS
DNA polymerase I (EC 2.7.7.7)
DNA polymerase III alpha subunit (EC 2.7.7.7)
DNA polymerase III beta subunit (EC 2.7.7.7)
DNA polymerase III chi subunit (EC 2.7.7.7)
DNA polymerase III delta prime subunit (EC 2.7.7.7)
DNA polymerase III delta subunit (EC 2.7.7.7)
DNA polymerase III epsilon subunit (EC 2.7.7.7)
DNA polymerase III subunits gamma and tau (EC 2.7.7.7)
DNA polymerase IV (EC 2.7.7.7)
DNA primase (EC 2.7.7.-)
DNA primase TraC (EC 2.7.7.-)
DNA recombination and repair protein RecF
DNA recombination and repair protein RecO
DNA recombination protein RmuC
DNA repair protein RadA
DNA repair protein RadC
DNA repair protein RecN
DNA replication protein DnaC

Nitric oxide reductase activation protein NorQ
Nitric oxide synthase oxygenase (EC 1.-.-.-)
Nitric-oxide reductase subunit B (EC 1.7.99.7)
Nitrite reductase accessory protein NirV
Nitrite-sensitive transcriptional repressor NsrR
Nitrogen regulation protein NR(I)
Nitrogen regulatory protein P-II
Nitrogen regulatory protein P-II, glnK
Nitrous oxide reductase maturation periplasmic protein NosX
Nitrous oxide reductase maturation protein NosD
Nitrous oxide reductase maturation protein, outer-membrane lipoprotein NosL
NnrU family protein, required for expression of nitric oxide and nitrite reductases (Nir and Nor)
Non-ribosomal peptide synthetase modules, siderophore biosynthesis
Nucleoside 5-triphosphatase RdgB (dHAPTP, dITP, XTP-specific) (EC 3.6.1.15)
O-succinylbenzoic acid--CoA ligase (EC 6.2.1.26)
Ornithine cyclodeaminase (EC 4.3.1.12)
Ornithine decarboxylase (EC 4.1.1.17)
Outer membrane ferripyoverdine receptor
Outer membrane receptor for ferric coprogen and ferric-rhodotorulic acid
PAB2272 methyltransferase homolog
Periplasmic [Fe] hydrogenase large subunit (EC 1.12.7.2)
periplasmic fructose-binding protein component of signal transduction system LevQ
periplasmic fructose-binding protein component of signal transduction system LevT
Periplasmic molybdate-binding domain
Phage capsid assembly
Phage capsid scaffolding protein
Phage capsid vertex
Phage capsid vertex assembly chaperone

D-beta-hydroxybutyrate dehydrogenase (EC 1.1.1.30)
D-erythrose-4-phosphate dehydrogenase (EC 1.2.1.72)
D-galactarate dehydratase (EC 4.2.1.42)
D-galactarate permease
D-galactonate transporter
D-glucarate permease
D-glycerate 2-kinase (EC 2.7.1.-)
D-glycero-beta-D-manno-heptose 7-phosphate kinase
D-hydantoinase (EC 3.5.2.2)
D-lactate dehydrogenase (EC 1.1.1.28)
D-Lactate dehydrogenase (EC 1.1.2.5)
D-Lactate dehydrogenase, cytochrome c-dependent (EC 1.1.2.4)
D-malic enzyme (EC 1.1.1.83)
D-mannonate oxidoreductase (EC 1.1.1.57)
D-mannose isomerase (EC 5.3.1.7)
D-serine dehydratase transcriptional activator
D-serine permease DsdX
D-serine/D-alanine/glycine transporter
D-serine/D-alanine/glycine transporter
D-xylose proton-symporter XylE
D-xylose transport ATP-binding protein XylG
D(-)-3-hydroxybutyrate oligomer hydrolase (EC 3.1.1.22)
Deacetylases, including yeast histone deacetylase and acetoin utilization protein
DEAD-box ATP-dependent RNA helicase CshA (EC 3.6.4.13)
Decaprenyl diphosphate synthase (EC 2.5.1.91)
Delta-1-pyrroline-5-carboxylate dehydrogenase (EC 1.5.1.12)
Deoxycytidine triphosphate deaminase (EC 3.5.4.13)
Deoxyguanosinetriphosphate triphosphohydrolase (EC 3.1.5.1)

DNA replication terminus site-binding protein
DNA topoisomerase I (EC 5.99.1.2)
DNA topoisomerase III (EC 5.99.1.2)
DNA topoisomerase III, Burkholderia type (EC 5.99.1.2)
DNA topoisomerase III, TraE-type (EC 5.99.1.2)
DNA transformation protein TfoX
DNA-3-methyladenine glycosylase (EC 3.2.2.20)
DNA-3-methyladenine glycosylase II (EC 3.2.2.21)
DNA-binding domain of ModE
DNA-binding heavy metal response regulator
DNA-binding protein Fis
DNA-binding protein HU
DNA-binding protein HU-beta
DNA-binding response regulator, LuxR family, near polyamine transporter
DNA-binding transcriptional activator of the allD operon
DNA-directed RNA polymerase alpha subunit (EC 2.7.7.6)
DNA-directed RNA polymerase beta' subunit (EC 2.7.7.6)
DnaJ-class molecular chaperone CbpA
Dolichol-phosphate mannosyltransferase (EC 2.4.1.83) in lipid-linked oligosaccharide synthesis cluster
dTDP-4-dehydrorhamnose 3,5-epimerase (EC 5.1.3.13)
dTDP-4-dehydrorhamnose reductase (EC 1.1.1.133)
dTDP-glucose 4,6-dehydratase (EC 4.2.1.46)
dTDP-rhamnosyl transferase RfbF (EC 2.-.-.-)
DUF1022 domain-containing protein
DUF1537 linked to pdxA
Duplicated ATPase component BL0693 of energizing module of predicted ECF transporter
Duplicated ATPase component CbrU of energizing module of predicted cobalamin

Phage capsid vertex protein
Phenazine biosynthesis protein PhzD
Phenazine biosynthesis protein PhzG
PhnH protein
Phosphate acetyltransferase (EC 2.3.1.8)
Phosphatidylglycerophosphatase B (EC 3.1.3.27)
Phosphatidylserine decarboxylase (EC 4.1.1.65)
Phosphogluconate dehydratase (EC 4.2.1.12)
Phosphoheptose isomerase 2 (EC 5.3.1.-)
Phosphonoacetate hydrolase (EC 3.11.1.2)
Phosphoribosyl-dephospho-CoA transferase (EC 2.7.7.-)
Phosphotransferase system for xylose-containing disaccharide, EIIA component
Photosystem I-type photosynthetic reaction center
Phycobilisome core component
Phycobilisome core-membrane linker polypeptide
Phycobilisome degradation protein NblA
Phycobilisome rod linker polypeptide, phycocyanin-associated
Phycobilisome rod linker polypeptide, phycoerythrocyanin-associated
Phycobilisome rod-core linker polypeptide, phycocyanin-associated
Phycobilisome small core linker polypeptide
Phycobilisome small rod capping linker polypeptide, phycocyanin-associated
Phycocyanin alpha chain
Phycocyanin alpha phycocyanobilin lyase related protein NblB
Phycocyanobilin lyase alpha subunit
Phycocyanobilin lyase beta subunit
Phycoerythrin alpha chain
Phycoerythrin beta chain

Deoxyhypusine synthase (EC 2.5.1.46)
Deoxyribodipyrimidine photolyase (EC 4.1.99.3)
Deoxyribonucleoside regulator DeoR (transcriptional repressor)
Deoxyribose operon repressor, DeoR family
Deoxyribose-phosphate aldolase (EC 4.1.2.4)
Dephospho-CoA kinase (EC 2.7.1.24)
Dethiobiotin synthetase (EC 6.3.3.3)
Di/tripeptide permease DtpA
Di/tripeptide permease YjdL
Diacylglycerol kinase (EC 2.7.1.107)
Diaminobutyrate-pyruvate aminotransferase (EC 2.6.1.46)
Diaminohydroxyphosphoribosylaminopyrimidine deaminase (EC 3.5.4.26)
Diaminopimelate decarboxylase (EC 4.1.1.20)
Diaminopimelate epimerase (EC 5.1.1.7)
dicarboxylic acid transporter PcaT
diglucosyldiacylglycerol synthase (LTA membrane anchor synthesis)
diguanylate cyclase/phosphodiesterase (GGDEF & EAL domains) with PAS/PAC sensor(s)
Dihydrofolate reductase (EC 1.5.1.3)
Dihydrofolate synthase (EC 6.3.2.12)
Dihydrolipoamide acetyltransferase component (E2) of acetoin dehydrogenase complex (EC 2.3.1.-)
Dihydrolipoamide acetyltransferase component of pyruvate dehydrogenase complex (EC 2.3.1.12)
Dihydrolipoamide acyltransferase component of branched-chain alpha-keto acid dehydrogenase complex (EC 2.3.1.168)
Dihydrolipoamide dehydrogenase (EC 1.8.1.4)
Dihydrolipoamide dehydrogenase of 2-oxoglutarate dehydrogenase (EC 1.8.1.4)
Dihydrolipoamide dehydrogenase of branched-chain alpha-keto acid dehydrogenase (EC 1.8.1.4)
Dihydrolipoamide dehydrogenase of pyruvate dehydrogenase complex (EC 1.8.1.4)
Dihydrolipoamide succinyltransferase component (E2) of 2-oxoglutarate dehydrogenase

ECF transporter
ElaA protein
Electron transfer flavoprotein-ubiquinone oxidoreductase (EC 1.5.5.1)
Electron transfer flavoprotein, alpha subunit
Electron transfer flavoprotein, beta subunit
Endo-1,4-beta-xylanase A precursor (EC 3.2.1.8)
Endonuclease III (EC 4.2.99.18)
Endonuclease IV (EC 3.1.21.2)
Endonuclease VIII
Enolase (EC 4.2.1.11)
Enoyl-[acyl-carrier-protein] reductase [FMN] (EC 1.3.1.9)
Enoyl-[acyl-carrier-protein] reductase [NADH] (EC 1.3.1.9)
Enoyl-[acyl-carrier-protein] reductase [NADPH] (EC 1.3.1.10)
Enoyl-CoA hydratase (EC 4.2.1.17)
Enoyl-CoA hydratase [valine degradation] (EC 4.2.1.17)
Enterobactin esterase
Enterobactin exporter EntS
Enterobactin synthetase component F, serine activating enzyme (EC 2.7.7.-)
Epi-inositol hydrolase (EC 3.7.1.-)
Erythritol kinase EryA (EC 2.7.1.27)
Ethanolamine ammonia-lyase heavy chain (EC 4.3.1.7)
Ethanolamine ammonia-lyase light chain (EC 4.3.1.7)
Evolved beta-D-galactosidase, beta subunit
Excinuclease ABC subunit A
Excinuclease ABC subunit A paralog in greater Bacteroides group
Excinuclease ABC subunit A paralog of unknown function
Excinuclease ABC subunit A, dimeric form
Excinuclease ABC subunit B
Excinuclease ABC subunit C

Phycoerythrin gamma chain linker polypeptide
Phycoerythrin linker protein CpeS homolog
Phycoerythrocyanin alpha chain
Phycoerythrocyanin beta chain
Plasmid partitioning protein ParA
Poly(glycerophosphate chain) D-alanine transfer protein DltD
Polysaccharide intercellular adhesin (PIA) biosynthesis protein IcaC
Portal protein [Bacteriophage A118]
Positive regulator of CheA protein activity (CheW)
Possible alpha-xyloside ABC transporter, ATP-binding component
Potassium-transporting ATPase C chain (EC 3.6.3.12) (TC 3.A.3.7.1)
Predicted alternative glutathione synthetase (EC 6.3.2.3)
Predicted b-glucoside-specific TonB-dependent outer membrane receptor
Predicted cobalt ABC transporter periplasmic component
Predicted cobalt transporter CbtA
Predicted flavin reductase RutF in novel pyrimidine catabolism pathway
Predicted glucose transporter in maltodextrin utilization gene cluster
Predicted L-arabinose ABC transport system, permease protein 2
Predicted L-rhamnose mutarotase
Predicted L-rhamnose permease RhaY
Predicted lactaldehyde dehydrogenase (EC 1.2.1.22)
Predicted mannitol permease
Predicted mannose transporter, GGP family
Predicted NAD regulator in Alphaproteobacteria
Predicted nicotinate-regulated transporter BH3254
Predicted rhamnogalacturonan lyase in rhamnose utilization cluster
Predicted rhamnogalacturonide-specific TRAP-type transporter, small transmembrane component RhiB
Predicted rhamnose oligosaccharide ABC transport system, permease

complex (EC 2.3.1.61)
Dihydroneopterin aldolase (EC 4.1.2.25)
Dihydroorotase (EC 3.5.2.3)
Dihydroorotate dehydrogenase (EC 1.3.3.1)
Dihydroorotate dehydrogenase electron transfer subunit (EC 1.3.3.1)
Dihydroorotate dehydrogenase, catalytic subunit (EC 1.3.3.1)
Dihydropteroate synthase (EC 2.5.1.15)
Dihydropyrimidinase (EC 3.5.2.2)
Dihydropyrimidine dehydrogenase [NADP+] (EC 1.3.1.2)
Dihydroxy-acid dehydratase (EC 4.2.1.9)
Dimethylhistidine N-methyltransferase (EC 2.1.1.44)
DinG family ATP-dependent helicase CPE1197
DinG family ATP-dependent helicase YoaA
Dipeptide transport ATP-binding protein DppD (TC 3.A.1.5.2)
Dipeptide transport ATP-binding protein DppF (TC 3.A.1.5.2)
Dipeptide transport system permease protein DppB (TC 3.A.1.5.2)
Dipeptide transport system permease protein DppC (TC 3.A.1.5.2)
Dipeptide-binding ABC transporter, periplasmic substrate-binding component (TC 3.A.1.5.2)
Dipeptidyl carboxypeptidase Dcp (EC 3.4.15.5)
DJ-1/YajL/PfpI superfamily, includes chaperone protein YajL (former ThiJ), parkinsonism-associated protein DJ-1, peptidases PfpI, Hsp31
DNA gyrase subunit A (EC 5.99.1.3)
DNA gyrase subunit B (EC 5.99.1.3)
DNA helicase IV
DNA ligase (EC 6.5.1.2)
DNA mismatch repair protein MutL
DNA mismatch repair protein MutS
DNA polymerase I (EC 2.7.7.7)
DNA polymerase II (EC 2.7.7.7)
DNA polymerase III alpha subunit (EC 2.7.7.7)

Exodeoxyribonuclease III (EC 3.1.11.2)
Exodeoxyribonuclease V alpha chain (EC 3.1.11.5)
Exodeoxyribonuclease V beta chain (EC 3.1.11.5)
Exodeoxyribonuclease V gamma chain (EC 3.1.11.5)
Exodeoxyribonuclease VII large subunit (EC 3.1.11.6)
Exodeoxyribonuclease VII small subunit (EC 3.1.11.6)
Exoenzymes regulatory protein AepA in lipid-linked oligosaccharide synthesis cluster
Exonuclease SbcC
Exopolyphosphatase (EC 3.6.1.11)
Exoribonuclease II (EC 3.1.13.1)
FAD pyrophosphatase (EC 3.6.1.18)
Fatty acid desaturase (EC 1.14.19.1)
Fe-S oxidoreductase-like protein in Rubrerythrin cluster
Fe-S-cluster-containing hydrogenase components 1
Ferredoxin
Ferredoxin reductase
Ferredoxin--NADP(+) reductase (EC 1.18.1.2)
Ferredoxin--nitrite reductase (EC 1.7.7.1)
Ferredoxin-dependent glutamate synthase (EC 1.4.7.1)
Ferredoxin-type protein NapG (periplasmic nitrate reductase)
Ferredoxin, 2Fe-2S
Ferric enterobactin transport system permease protein FepD (TC 3.A.1.14.2)
Ferric iron ABC transporter, permease protein
Ferric siderophore transport system, biopolymer transport protein ExbB
Ferric siderophore transport system, periplasmic binding protein TonB

component
Predicted rhamnose oligosaccharide ABC transport system, permease component 2
Predicted rhamnose oligosaccharide ABC transport system, substrate-binding component
Predicted signal transduction protein
Predicted sodium-dependent mannose transporter
Predicted transcriptional regulator of 4-carboxymuconolactone decarboxylase, Rrf2 family
Predicted transcriptional regulator of cysteine synthase, Rrf2 family
Predicted transcriptional regulator of leucine degradation pathway, MerR family
Predicted transcriptional regulator of N-Acetylglucosamine utilization, LacI family
Predicted transcriptional regulator of NADH dehydrogenase, Rrf2 family
Predicted transcriptional regulator of sulfate adenylyltransferase, Rrf2 family
Predicted transcriptional regulator of sulfate transport, Rrf2 family
Predicted trehalose permease, MFS family
Prephenate and/or arogenate dehydrogenase (unknown specificity) (EC 1.3.1.12)(EC 1.3.1.43)
Prephenate dehydrogenase (EC 1.3.1.12)
Probable acyl-CoA dehydrogenase (EC 1.3.99.3)
Probable electron transfer flavoprotein-quinone oxidoreductase FixC (EC 1.5.5.-)
Probable L-ascorbate-6-phosphate lactonase UlaG (EC 3.1.1.-) (L-ascorbate utilization protein G)
Probable transcriptional regulator of trehalose utilization, LuxR family
Probable two-component sensor, near polyamine transporter
Proline iminopeptidase (EC 3.4.11.5)
Proline racemase (EC 5.1.1.4)
Proline-specific permease proY
Propanediol dehydratase small subunit (EC 4.2.1.28)
Propionate--CoA ligase (EC 6.2.1.17)
Protein acetyltransferase

DNA polymerase III beta subunit (EC 2.7.7.7)
DNA polymerase III chi subunit (EC 2.7.7.7)
DNA polymerase III delta prime subunit (EC 2.7.7.7)
DNA polymerase III delta subunit (EC 2.7.7.7)
DNA polymerase III epsilon subunit (EC 2.7.7.7)
DNA polymerase III subunits gamma and tau (EC 2.7.7.7)
DNA polymerase III theta subunit (EC 2.7.7.7)
DNA polymerase IV (EC 2.7.7.7)
DNA polymerase-like protein PA0670
DNA primase (EC 2.7.7.-)
DNA primase (EC 2.7.7.-), phage-associated
DNA recombination and repair protein RecF
DNA recombination and repair protein RecO
DNA recombination protein RmuC
DNA recombination-dependent growth factor C
DNA repair protein RadA
DNA repair protein RadC
DNA repair protein RecN
DNA replication protein DnaC
DNA replication terminus site-binding protein
DNA topoisomerase I (EC 5.99.1.2)
DNA topoisomerase III (EC 5.99.1.2)
DNA topoisomerase III, Burkholderia type (EC 5.99.1.2)
DNA topoisomerase VI subunit A (EC 5.99.1.3)
DNA topoisomerase VI subunit B (EC 5.99.1.3)

Ferric uptake regulation protein FUR
Ferrichrome transport ATP-binding protein FhuC (TC 3.A.1.14.3)
Ferrichrome-iron receptor
Ferrochelatase, protoheme ferro-lyase (EC 4.99.1.1)
Ferrous iron transport periplasmic protein EfeO, contains peptidase-M75 domain and (frequently) cupredoxin-like domain
Ferrous iron transport protein B
FIG002571: 4-hydroxybenzoyl-CoA thioesterase domain protein
FIG003492: Threonine dehydrogenase and related Zn-dependent dehydrogenases
FIG005121: SAM-dependent methyltransferase (EC 2.1.1.-)
FIG00637966: hypothetical protein
FIG007491: hypothetical protein YeeN
FIG00820327: hypothetical protein
FIG017823: ATPase, MoxR family
FIG021862: membrane protein, exporter
FIG033889: YebC paralog in Betaproteobacteria
FIG039061: hypothetical protein related to heme utilization
FIG049476: HIT family protein
FIG138315: Putative alpha helix protein
FIG138576: 3-oxoacyl-[ACP] synthase (EC 2.3.1.41)
FIG143828: Hypothetical protein YbgA
Fimbrial protein YadM
FKBP-type peptidyl-prolyl cis-trans isomerase SlyD (EC 5.2.1.8)
Flagellar assembly protein FliH
Flagellar basal-body P-ring formation protein FlgA
Protein clustered with ethanolamine utilization
Protein GerPA, required for proper assembly of spore coat, mutations lead to super-dormant spore
Protein GerPB, required for proper assembly of spore coat, mutations lead to super-dormant spore
Protein GerPC, required for proper assembly of spore coat, mutations lead to super-dormant spore
Protein GerPF, required for proper assembly of spore coat, mutations lead to super-dormant spore
Protein gp13 [Bacteriophage A118]
Protein gp32 [Listeria phage 2389]
Protein gp34 [Bacteriophage A118]
Protein gp51 [Bacteriophage A118]
Protein gp68 [Bacteriophage A118]
Protein gp7 [Bacteriophage A118]
Protein gp9 [Bacteriophage A118]
Protein involved in meta-pathway of phenol degradation
Protein similar to glutamate synthase [NADPH] small chain, clustered with sulfite reductase
Protein-L-isoaspartate O-methyltransferase (EC 2.1.1.77)
Pseudaminic acid cytidylyltransferase (EC 2.7.7.43)
Pseudaminic acid synthase (EC 4.1.3.-)
PTS system, fructose- and mannose-inducible IIA component (EC 2.7.1.69)
PTS system, fructose- and mannose-inducible IID component (EC 2.7.1.69)
PTS system, glucitol/sorbitol-specific IIB component and second of two IIC components (EC 2.7.1.69)
PTS system, glucose-specific IIB component (EC 2.7.1.69)
PTS system, mannose-specific IIB component (EC 2.7.1.69)
PTS system, sorbose-specific IIB component (EC 2.7.1.69)
Pup ligase PafA' paralog, possible component of postulated heterodimer PafA-PafA'
DNA-3-methyladenine glycosylase (EC 3.2.2.20)
DNA-3-methyladenine glycosylase II (EC 3.2.2.21)
DNA-binding domain of ModE
DNA-binding protein HU-beta
DNA-binding response regulator, LuxR family, near polyamine transporter
DNA-directed RNA polymerase alpha subunit (EC 2.7.7.6)
DNA-directed RNA polymerase beta' subunit (EC 2.7.7.6)
DNA-directed RNA polymerase omega subunit (EC 2.7.7.6)
DnaJ-class molecular chaperone CbpA
Dolichol-phosphate mannosyltransferase (EC 2.4.1.83) in lipid-linked oligosaccharide synthesis cluster
dTDP-4-dehydrorhamnose 3,5-epimerase (EC 5.1.3.13)
dTDP-4-dehydrorhamnose reductase (EC 1.1.1.133)
dTDP-glucose 4,6-dehydratase (EC 4.2.1.46)
DUF1022 domain-containing protein
DUF1537 linked to pdxA
Duplicated ATPase component CbrU of energizing module of predicted cobalamin ECF transporter
Duplicated ATPase component YkoD of energizing module of thiamin-regulated ECF transporter for HydroxyMethylPyrimidine
Ectoine hydroxylase (EC 1.17.-.-)
ElaA protein
Electron transfer flavoprotein-ubiquinone oxidoreductase (EC 1.5.5.1)
Electron transfer flavoprotein, alpha subunit
Electron transfer flavoprotein, alpha subunit FixB
Electron transfer flavoprotein, beta subunit
Electron transport complex protein RnfD

Flagellar basal-body rod modification protein FlgD
Flagellar basal-body rod protein FlgB
Flagellar basal-body rod protein FlgC
Flagellar basal-body rod protein FlgF
Flagellar basal-body rod protein FlgG
Flagellar biosynthesis protein FlhA
Flagellar biosynthesis protein FlhB
Flagellar biosynthesis protein FliC
Flagellar biosynthesis protein FliL
Flagellar biosynthesis protein FliP
Flagellar biosynthesis protein FliQ
Flagellar biosynthesis protein FliR
Flagellar biosynthesis protein FliS
Flagellar hook protein FlgE
Flagellar hook-associated protein FlgK
Flagellar hook-associated protein FlgL
Flagellar hook-associated protein FliD
Flagellar hook-basal body complex protein FliE
Flagellar L-ring protein FlgH
Flagellar M-ring protein FliF
Flagellar motor rotation protein MotA
Flagellar motor rotation protein MotB
Flagellar motor switch protein FliG
Flagellar motor switch protein FliM
Flagellar motor switch protein FliN
Flagellar P-ring protein FlgI
Flagellar protein FlgJ [peptidoglycan hydrolase] (EC 3.2.1.-)
Flagellar protein FliJ

Putative ABC transporter of substrate X, ATP-binding subunit
Putative ABC transporter of substrate X, permease subunit I
Putative arylsulfatase regulatory protein
Putative conjugative transposon mobilization protein BF0132
putative Cytochrome bd2, subunit II
Putative cytoplasmic protein clustered with trehalase
Putative dihydroxyacetone kinase (EC 2.7.1.29), dihydroxyacetone binding subunit
Putative exported protein clustered with Gamma-glutamyltranspeptidase
putative facilitator of salicylate uptake
Putative glucanase glgE (EC 3.2.1.-)
Putative GntR-family regulatory protein and aminotransferase near polyamine transporter
Putative heme iron utilization protein
Putative mobilization protein BF0133
Putative N-carbamyl-L-cysteine amidohydrolase
Putative OMR family iron-siderophore receptor precursor
Putative oxidoreductase in 4-hydroxyproline catabolic gene cluster
Putative peptidoglycan bound protein (LPXTG motif) Lmo0835 homolog
Putative peptidoglycan bound protein (LPXTG motif) Lmo0842 homolog
Putative peptidoglycan bound protein (LPXTG motif) Lmo1799 homolog
Putative peptidoglycan bound protein (LPXTG motif) Lmo2714 homolog
Putative peptidoglycan hydrolase YvbX, NOT involved in spore germination
Putative phthalate 4,5-dioxygenase oxygenase subunit (OhpA2)
Putative regulator of the mannose operon, ManO
Putative silver efflux pump
putative succinate dehydrogenase [membrane anchor subunit] (succinic dehydrogenase)
Putative succinate dehydrogenase cytochrome b subunit
Putative sugar nucleotidyltransferase
putative TolA function

Endonuclease III (EC 4.2.99.18)
Endonuclease IV (EC 3.1.21.2)
Endonuclease V (EC 3.1.21.7)
Enolase (EC 4.2.1.11)
Enoyl-[acyl-carrier-protein] reductase [FMN] (EC 1.3.1.9)
Enoyl-[acyl-carrier-protein] reductase [NADH] (EC 1.3.1.9)
Enoyl-[acyl-carrier-protein] reductase [NADPH] (EC 1.3.1.10)
Enoyl-CoA hydratase (EC 4.2.1.17)
Enoyl-CoA hydratase [valine degradation] (EC 4.2.1.17)
Enterobactin esterase
Enterobactin exporter EntS
Enterobactin synthetase component F, serine activating enzyme (EC 2.7.7.-)
Epi-inositol hydrolase (EC 3.7.1.-)
Erythritol kinase EryA (EC 2.7.1.27)
Erythritol phosphate dehydrogenase EryB
Ethanolamine ammonia-lyase heavy chain (EC 4.3.1.7)
Ethanolamine ammonia-lyase light chain (EC 4.3.1.7)
Ethylmalonyl-CoA epimerase
Evolved beta-D-galactosidase, alpha subunit
Excinuclease ABC subunit A
Excinuclease ABC subunit A paralog of unknown function
Excinuclease ABC subunit A, dimeric form
Excinuclease ABC subunit B
Excinuclease ABC subunit C
Excinuclease cho (excinuclease ABC alternative C subunit)
Exodeoxyribonuclease I (EC 3.1.11.1)
Exodeoxyribonuclease III (EC 3.1.11.2)
Exodeoxyribonuclease V alpha chain (EC 3.1.11.5)

Flagellar regulatory protein FleQ

Flagellar sensor histidine kinase FleS

Flagellar transcriptional activator FlhC

Flagellin protein FlaA

Flagellin protein FlaC

Flagellum-specific ATP synthase FliI

Flavodoxin 2

Flavodoxin reductases (ferredoxin-NADPH reductases) family 1

Flavohemoprotein (Hemoglobin-like protein) (Flavohemoglobin) (Nitric oxide dioxygenase) (EC 1.14.12.17)

Flp pilus assembly protein CpaD

Flp pilus assembly protein RcpC/CpaB

Flp pilus assembly protein TadB

Folate-dependent protein for Fe/S cluster synthesis/repair in oxidative stress

Foldase protein PrsA precursor (EC 5.2.1.8)

FolM Alternative dihydrofolate reductase 1

Folylpolyglutamate synthase (EC 6.3.2.17)

Formamidopyrimidine-DNA glycosylase (EC 3.2.2.23)

Formate dehydrogenase chain D (EC 1.2.1.2)

formate dehydrogenase formation protein FdhE

Formate dehydrogenase O alpha subunit (EC 1.2.1.2)

Formate dehydrogenase O beta subunit (EC 1.2.1.2)

Formate dehydrogenase O gamma subunit (EC 1.2.1.2)

Formate dehydrogenase-O, major subunit (EC 1.2.1.2)

Formate efflux transporter (TC 2.A.44 family)

Formate hydrogenlyase subunit 3

Formate hydrogenlyase subunit 3/Multisubunit Na+/H+ antiporter, MnhD subunit

Formate hydrogenlyase subunit 4

Formate--tetrahydrofolate ligase (EC 6.3.4.3)

Putative TolA protein

Putative TolA protein involved in Tol transport system in Chlamydia

Putative Ton-B dependent hemine receptor

Putative TonB-dependent heme receptor HasR

PutR, transcriptional activator of PutA and PutP

Putrescine ABC transporter putrescine-binding protein PotF (TC 3.A.1.11.2)

Putrescine carbamoyltransferase (EC 2.1.3.6)

Putrescine transport ATP-binding protein PotA (TC 3.A.1.11.1)

Putrescine transport ATP-binding protein PotG (TC 3.A.1.11.2)

Putrescine transport system permease protein PotH (TC 3.A.1.11.2)

Putrescine transport system permease protein PotI (TC 3.A.1.11.2)

Putrescine-ornithine antiporter

Putrescine/proton symporter, putrescine/ornithine antiporter PotE

pyrrolysine-containing

Pyruvate dehydrogenase E1 component beta subunit (EC 1.2.4.1)

Pyruvate formate-lyase (EC 2.3.1.54)

QscR quorum-sensing control repressor

Quinolinate phosphoribosyltransferase [decarboxylating] (EC 2.4.2.19)

Quinone-reactive Ni/Fe hydrogenase, cytochrome b subunit

Radical SAM family enzyme, similar to coproporphyrinogen III oxidase, oxygen-independent, clustered with nucleoside-triphosphatase RdgB

Rare lipoprotein A precursor

Regulatory protein RsaL

RelE antibacterial toxin protein

Repressor (cro-like) [Bacteriophage A118]

Response regulator of zinc sigma-54-dependent two-component system

rhamnogalacturonan acetylesterase

Rhamnogalacturonide transporter RhiT

Rhamnolipids biosynthesis 3-oxoacyl-[acyl-carrier-protein] reductase RhlG

Exodeoxyribonuclease V beta chain (EC 3.1.11.5)

Exodeoxyribonuclease V gamma chain (EC 3.1.11.5)

Exodeoxyribonuclease VII large subunit (EC 3.1.11.6)

Exodeoxyribonuclease VII small subunit (EC 3.1.11.6)

Exoenzymes regulatory protein AepA in lipid-linked oligosaccharide synthesis cluster

Exonuclease SbcC

Exonuclease SbcD

Exopolyphosphatase (EC 3.6.1.11)

Exoribonuclease II (EC 3.1.13.1)

F420-dependent N(5),N(10)-methylenetetrahydromethanopterin reductase (EC 1.5.99.11)

Fatty acid desaturase (EC 1.14.19.1)

Fe-S-cluster-containing hydrogenase components 1

Ferredoxin

Ferredoxin reductase

Ferredoxin--NADP(+) reductase (EC 1.18.1.2)

Ferredoxin--nitrite reductase (EC 1.7.7.1)

Ferredoxin-dependent glutamate synthase (EC 1.4.7.1)

Ferredoxin-like protein YgcO

Ferredoxin-type protein NapF (periplasmic nitrate reductase)

Ferredoxin, 2Fe-2S

Ferric enterobactin transport ATP-binding protein FepC (TC 3.A.1.14.2)

Ferric enterobactin uptake protein FepE

Ferric enterobactin-binding periplasmic protein FepB (TC 3.A.1.14.2)

Ferric hydroxamate ABC transporter (TC 3.A.1.14.3), ATP-binding protein FhuC

Ferric hydroxamate ABC transporter (TC 3.A.1.14.3), periplasmic substrate binding protein FhuD

Ferric hydroxamate ABC transporter (TC 3.A.1.14.3), permease component FhuB

Ferric hydroxamate outer membrane receptor FhuA

Ferric iron ABC transporter, ATP-binding protein

Formiminoglutamic iminohydrolase (EC 3.5.3.13)
Forms the bulk of type IV secretion complex that spans outer membrane and periplasm (VirB9)
Formyltetrahydrofolate deformylase (EC 3.5.1.10)
Frataxin homolog CyaY, facilitates iron supply for heme A synthesis or Fe-S cluster assembly
Fructokinase (EC 2.7.1.4)
Fructose ABC transporter, ATP-binding component FrcA
Fructose-1,6-bisphosphatase, GlpX type (EC 3.1.3.11)
Fructose-1,6-bisphosphatase, type I (EC 3.1.3.11)
Fructose-bisphosphate aldolase class I (EC 4.1.2.13)
Fructose-bisphosphate aldolase class II (EC 4.1.2.13)
Fumarate and nitrate reduction regulatory protein
Fumarate hydratase class I (EC 4.2.1.2)
Fumarate hydratase class I, aerobic (EC 4.2.1.2)
Fumarate hydratase class II (EC 4.2.1.2)
Fumarate reductase flavoprotein subunit (EC 1.3.99.1)
Fumarylacetoacetase (EC 3.7.1.2)
Fumarylacetoacetate (FAA) hydrolase (EC 4.1.1.68)
Fumarylacetoacetate hydrolase family protein
G:T/U mismatch-specific uracil/thymine DNA-glycosylase
Galactitol utilization operon repressor
Galactokinase (EC 2.7.1.6)
Galactonate dehydratase (EC 4.2.1.6)
Galactosamine-6-phosphate isomerase (EC 5.3.1.-)
Galactose-6-phosphate isomerase, LacB subunit (EC 5.3.1.26)
Galactose/methyl galactoside ABC transport system, ATP-binding protein MglA (EC 3.6.3.17)
Galactoside O-acetyltransferase (EC 2.3.1.18)

(EC 1.1.1.100)
Rhamnulokinase RhaK in alpha-proteobacteria (EC 2.7.1.5)
RhlA, 3-(3-hydroxyalkanoyloxy)alkanoic acids (HAAs) synthase
RhlB, TDP-rhamnosyltransferase 1 (EC 2.4.1.-)
RhlC, TDP-rhamnosyltransferase 2 (EC 2.4.1.-)
Riboflavin synthase archaeal (EC 2.5.1.9)
Riboflavin transporter PnuX
Ribonuclease I precursor (EC 3.1.27.6)
Ribose 5-phosphate isomerase A (EC 5.3.1.6)
Ribosome-associated heat shock protein implicated in the recycling of the 50S subunit (S4 paralog)
Ribosome-binding factor A
Ribulose bisphosphate carboxylase large chain (EC 4.1.1.39)
Ribulose bisphosphate carboxylase small chain (EC 4.1.1.39)
Ribulose-phosphate 3-epimerase (EC 5.1.3.1)
RNA polymerase sigma factor RpoD
RNA polymerase sigma factor RpoH
RNA polymerase sigma-54 factor RpoN
RNA polymerase sporulation specific sigma factor SigE
RND efflux system, outer membrane lipoprotein, NodT family
Rod shape-determining protein MreC
Rod shape-determining protein MreD
Rrf2 family transcriptional regulator
Rrf2 family transcriptional regulator, group III
S-adenosylhomocysteine nucleosidase (EC 3.2.2.9)
S-adenosylmethionine decarboxylase proenzyme (EC 4.1.1.50), eukaryotic
S-adenosylmethionine decarboxylase proenzyme (EC 4.1.1.50), prokaryotic class 1A
S-adenosylmethionine decarboxylase proenzyme (EC 4.1.1.50), prokaryotic class 1B

Ferric iron ABC transporter, iron-binding protein
Ferric iron ABC transporter, permease protein
Ferric reductase (1.6.99.14)
Ferric siderophore transport system, periplasmic binding protein TonB
Ferric uptake regulation protein FUR
Ferrochelatase, protoheme ferro-lyase (EC 4.99.1.1)
Ferrous iron transport periplasmic protein EfeO, contains peptidase-M75 domain and (frequently) cupredoxin-like domain
Ferrous iron transport peroxidase EfeB
Ferrous iron transport protein A
Ferrous iron transport protein B
FIG000605: protein co-occurring with transport systems (COG1739)
FIG001341: Probable Fe(2+)-trafficking protein YggX
FIG002003: Protein YdjA
FIG002571: 4-hydroxybenzoyl-CoA thioesterase domain protein
FIG003492: Threonine dehydrogenase and related Zn-dependent dehydrogenases
FIG005121: SAM-dependent methyltransferase (EC 2.1.1.-)
FIG00637966: hypothetical protein
FIG017823: ATPase, MoxR family
FIG018171: hypothetical protein of Cupin superfamily
FIG018329: 1-acyl-sn-glycerol-3-phosphate acyltransferase
FIG027190: Putative transmembrane protein
FIG043197: Inositol monophosphatase family protein
FIG138315: Putative alpha helix protein
FIG138576: 3-oxoacyl-[ACP] synthase (EC 2.3.1.41)
FIG146085: 3'-to-5' oligoribonuclease A, Bacillus type
FIGfam014588: Predicted regulator of CFA/I fimbriae

Gamma-aminobutyrate:alpha-ketoglutarate aminotransferase (EC 2.6.1.19)
Gamma-glutamyl phosphate reductase (EC 1.2.1.41)
Gamma-glutamyl-aminobutyraldehyde dehydrogenase (EC 1.2.1.-)
Gamma-glutamyl-GABA hydrolase (EC 3.5.1.94)
Gamma-glutamyl-putrescine synthetase (EC 6.3.1.11)
Gamma-glutamyltranspeptidase (EC 2.3.2.2)
Gamma-glutamyltranspeptidase PgsD/CapD (EC 2.3.2.2), catalyses PGA anchorage to peptidoglycan
GbcA Glycine betaine demethylase subunit A
GDP-L-fucose synthetase (EC 1.1.1.271)
GDP-mannose 4,6-dehydratase (EC 4.2.1.47)
GDP-mannose pyrophosphatase YffH
General secretion pathway protein C
General secretion pathway protein F
General secretion pathway protein L
General secretion pathway protein M
Gentisate 1,2-dioxygenase (EC 1.13.11.4)
Geranyl-CoA carboxylase biotin-containing subunit (EC 6.4.1.5)
Geranylgeranyl diphosphate reductase (EC 1.3.1.83)
Geranylgeranyl pyrophosphate synthetase (EC 2.5.1.29)
GlpG protein (membrane protein of glp regulon)
Glucans biosynthesis glucosyltransferase H (EC 2.4.1.-)
Glucarate dehydratase (EC 4.2.1.40)
Glucoamylase (EC 3.2.1.3)
Glucokinase (EC 2.7.1.2)
Gluconate dehydratase (EC 4.2.1.39)
Gluconate permease
Gluconate transporter family protein
Gluconate utilization system Gnt-I transcriptional repressor

S-ribosylhomocysteine lyase (EC 4.4.1.21)
Sarcosine N-methyltransferase
Secreted protein, suppressor for copper-sensitivity ScsC
Sensor histidine kinase colocalized with HrtAB transporter
Sensor histidine protein kinase UhpB, glucose-6-phosphate specific (EC 2.7.13.3)
Sensor protein of zinc sigma-54-dependent two-component system
Sensor protein PhoQ (EC 2.7.13.3)
Serine hydroxymethyltransferase (EC 2.1.2.1)
Serine--glyoxylate aminotransferase (EC 2.6.1.45)
Shikimate 5-dehydrogenase (EC 1.1.1.25)
Shikimate 5-dehydrogenase I gamma (EC 1.1.1.25)
Shikimate kinase I (EC 2.7.1.71)
Sialic acid-regulated TonB-dependent outer membrane receptor
Signal transduction histidine kinase CheA (EC 2.7.3.-)
Similar to eukaryotic Peptidyl prolyl 4-hydroxylase, alpha subunit (EC 1.14.11.2)
Single-stranded DNA-binding protein
SMT0608 replicon stabilization toxin
SMT0609 replicon stabilization protein (antitoxin to SMT0608)
Sorbitol operon transcription regulator
Sorbitol-6-phosphate 2-dehydrogenase (EC 1.1.1.140)
Sortase
Spermidine Putrescine ABC transporter permease component PotB (TC 3.A.1.11.1)
Spermidine Putrescine ABC transporter permease component potC (TC_3.A.1.11.1)
Spermidine synthase (EC 2.5.1.16)
Spore cortex-lytic enzyme CwlJ
Spore cortex-lytic enzyme, lytic transglycosylase SleB
Spore germination protein GerD
Spore germination protein GerKC

Fimbrial protein YadM
FKBP-type peptidyl-prolyl cis-trans isomerase FkpA precursor (EC 5.2.1.8)
FKBP-type peptidyl-prolyl cis-trans isomerase SlyD (EC 5.2.1.8)
Flagellar basal-body P-ring formation protein FlgA
Flagellar basal-body rod modification protein FlgD
Flagellar basal-body rod protein FlgB
Flagellar basal-body rod protein FlgC
Flagellar basal-body rod protein FlgF
Flagellar basal-body rod protein FlgG
Flagellar biosynthesis protein FlhA
Flagellar biosynthesis protein FlhB
Flagellar biosynthesis protein FlhF
Flagellar biosynthesis protein FliL
Flagellar biosynthesis protein FliP
Flagellar biosynthesis protein FliQ
Flagellar biosynthesis protein FliR
Flagellar biosynthesis protein FliS
Flagellar biosynthesis protein FliZ
Flagellar hook protein FlgE
Flagellar hook-associated protein FlgK
Flagellar hook-associated protein FlgL
Flagellar hook-associated protein FliD
Flagellar hook-basal body complex protein FliE
Flagellar hook-length control protein FliK
Flagellar L-ring protein FlgH
Flagellar M-ring protein FliF
Flagellar motor rotation protein MotA
Flagellar motor rotation protein MotB

Gluconokinase (EC 2.7.1.12)
Gluconolactonase (EC 3.1.1.17)
Glucosamine--fructose-6-phosphate aminotransferase [isomerizing] (EC 2.6.1.16)
Glucosamine-6-phosphate deaminase (EC 3.5.99.6)
Glucosamine-6-phosphate deaminase [isomerizing], alternative (EC 3.5.99.6)
Glucose 1-dehydrogenase (EC 1.1.1.47)
Glucose dehydrogenase (EC 1.1.99.10), membrane-bound, flavoprotein
Glucose dehydrogenase, PQQ-dependent (EC 1.1.5.2)
Glucose-1-phosphate adenylyltransferase (EC 2.7.7.27)
Glucose-1-phosphate thymidylyltransferase (EC 2.7.7.24)
Glucose-6-phosphate 1-dehydrogenase (EC 1.1.1.49)
Glucose-6-phosphate isomerase (EC 5.3.1.9)
Glucose-methanol-choline (GMC) oxidoreductase:NAD binding site
Glucuronide transporter UidB
Glutamate 5-kinase (EC 2.7.2.11)
Glutamate N-acetyltransferase (EC 2.3.1.35)
Glutamate N-acetyltransferase (EC 2.3.1.35) / N-acetylglutamate synthase (EC 2.3.1.1)
Glutamate racemase (EC 5.1.1.3)
Glutamate synthase [NADPH] large chain (EC 1.4.1.13)
Glutamate synthase [NADPH] small chain (EC 1.4.1.13)
Glutamate--cysteine ligase (EC 6.3.2.2)
Glutamate--cysteine ligase (EC 6.3.2.2), divergent, of Alpha- and Beta-proteobacteria type
Glutamate-1-semialdehyde aminotransferase (EC 5.4.3.8)
Glutamate-ammonia-ligase adenylyltransferase (EC 2.7.7.42)
Glutaminase (EC 3.5.1.2)
glutamine synthetase family protein
Glutamine synthetase type I (EC 6.3.1.2)
Glutamine synthetase type II, eukaryotic (EC 6.3.1.2)
Glutamine synthetase type III, GlnN (EC 6.3.1.2)

Spore germination protein GerLA
Spore germination protein GerLC
Spore germination protein GerQC
Spore germination protein GerYA
Spore protein YdhD, not involved in spore germination
Sporulation initiation phosphotransferase (Spo0F)
SSU ribosomal protein MRP1, mitochondrial
SSU ribosomal protein MRP13, mitochondrial
SSU ribosomal protein PET123, mitochondrial
SSU ribosomal protein RSM22, mitochondrial
SSU ribosomal protein RSM24, mitochondrial
SSU ribosomal protein RSM28, mitochondrial
SSU ribosomal protein S10p (S20e), mitochondrial
SSU ribosomal protein S11e (S17p)
SSU ribosomal protein S11p (S14e), mitochondrial
SSU ribosomal protein S12p (S23e), mitochondrial
SSU ribosomal protein S13e (S15p)
SSU ribosomal protein S13p (S18e), chloroplast
SSU ribosomal protein S13p (S18e), mitochondrial
SSU ribosomal protein S14p (S29e), mitochondrial
SSU ribosomal protein S15Ae (S8p)
SSU ribosomal protein S15p (S13e)
SSU ribosomal protein S15p (S13e)
SSU ribosomal protein S15p (S13e), mitochondrial
SSU ribosomal protein S16e (S9p)
SSU ribosomal protein S16p, mitochondrial
SSU ribosomal protein S17e
SSU ribosomal protein S17p (S11e), chloroplast
SSU ribosomal protein S17p (S11e), mitochondrial

Flagellar motor switch protein FliG
Flagellar motor switch protein FliM
Flagellar motor switch protein FliN
Flagellar P-ring protein FlgI
Flagellar protein FlbB
Flagellar protein FlgJ [peptidoglycan hydrolase] (EC 3.2.1.-)
Flagellar protein FliJ
Flagellar regulatory protein FleQ
Flagellar transcriptional activator FlhC
Flagellin protein FlaA
Flagellum-specific ATP synthase FliI
Flavodoxin 2
Flavodoxin reductases (ferredoxin-NADPH reductases) family 1
Flavohemoprotein (Hemoglobin-like protein) (Flavohemoglobin) (Nitric oxide dioxygenase) (EC 1.14.12.17)
Flp pilus assembly protein CpaD
Flp pilus assembly protein RcpC/CpaB
Flp pilus assembly protein TadB
FMN adenylyltransferase (EC 2.7.7.2)
FMN reductase (EC 1.5.1.29)
Folate-dependent protein for Fe/S cluster synthesis/repair in oxidative stress
Foldase protein PrsA precursor (EC 5.2.1.8)
FolM Alternative dihydrofolate reductase 1
Folylpolyglutamate synthase (EC 6.3.2.17)
Formamidopyrimidine-DNA glycosylase (EC 3.2.2.23)
Formate dehydrogenase -O, gamma subunit (EC 1.2.1.2)
Formate dehydrogenase chain D (EC 1.2.1.2)
formate dehydrogenase formation protein FdhE
Formate dehydrogenase N alpha subunit (EC 1.2.1.2)
Formate dehydrogenase N gamma subunit (EC 1.2.1.2)

Glutaminyl-tRNA synthetase (EC 6.1.1.18)
glutamyl-Q-tRNA synthetase
Glutamyl-tRNA reductase (EC 1.2.1.70)
Glutamyl-tRNA synthetase (EC 6.1.1.17)
Glutamyl-tRNA(Gln) amidotransferase subunit A (EC 6.3.5.7)
Glutamyl-tRNA(Gln) amidotransferase subunit B (EC 6.3.5.7)
Glutamyl-tRNA(Gln) synthetase (EC 6.1.1.24)
Glutaredoxin 3 (Grx2)
Glutaredoxin-related protein
Glutathione peroxidase (EC 1.11.1.9)
Glutathione reductase (EC 1.8.1.7)
Glutathione S-transferase (EC 2.5.1.18)
Glutathione S-transferase family protein
Glutathione S-transferase, omega (EC 2.5.1.18)
Glutathione S-transferase, unnamed subgroup (EC 2.5.1.18)
Glutathione synthetase (EC 6.3.2.3)
Glutathione-dependent formaldehyde-activating enzyme (EC 4.4.1.22)
Glutathione-regulated potassium-efflux system ATP-binding protein
Glutathione-regulated potassium-efflux system protein KefB
Glutathione-regulated potassium-efflux system protein KefC
Glycerate kinase (EC 2.7.1.31)
Glycerol dehydrogenase (EC 1.1.1.6)
Glycerol kinase (EC 2.7.1.30)
Glycerol uptake facilitator protein
Glycerol uptake operon antiterminator regulatory protein
Glycerol-1-phosphate dehydrogenase [NAD(P)] (EC 1.1.1.261)
Glycerol-3-phosphate ABC transporter, ATP-binding protein UgpC (TC 3.A.1.1.3)
Glycerol-3-phosphate ABC transporter, periplasmic glycerol-3-phosphate-binding protein (TC 3.A.1.1.3)
Glycerol-3-phosphate ABC transporter, permease protein UgpA (TC 3.A.1.1.3)
Glycerol-3-phosphate ABC transporter, permease protein UgpE (TC 3.A.1.1.3)

SSU ribosomal protein S18e (S13p)
SSU ribosomal protein S18p, mitochondrial
SSU ribosomal protein S19e
SSU ribosomal protein S19p (S15e), mitochondrial
SSU ribosomal protein S20e (S10p)
SSU ribosomal protein S21p, mitochondrial
SSU ribosomal protein S23e (S12p)
SSU ribosomal protein S23mt, mitochondrial
SSU ribosomal protein S24mt, mitochondrial
SSU ribosomal protein S25mt, mitochondrial
SSU ribosomal protein S26e
SSU ribosomal protein S26mt, mitochondrial
SSU ribosomal protein S27e
SSU ribosomal protein S27mt, mitochondrial
SSU ribosomal protein S28e
SSU ribosomal protein S28mt, mitochondrial
SSU ribosomal protein S29e (S14p)
SSU ribosomal protein S29mt, mitochondrial
SSU ribosomal protein S2p (SAe)
SSU ribosomal protein S2p (SAe), mitochondrial
SSU ribosomal protein S30e
SSU ribosomal protein S30mt, mitochondrial
SSU ribosomal protein S31mt, mitochondrial
SSU ribosomal protein S33mt, mitochondrial
SSU ribosomal protein S34mt, mitochondrial
SSU ribosomal protein S3Ae
SSU ribosomal protein S3e (S3p)
SSU ribosomal protein S3p (S3e), mitochondrial
SSU ribosomal protein S4e
SSU ribosomal protein S4p (S9e)

Formate dehydrogenase O alpha subunit (EC 1.2.1.2)
Formate dehydrogenase O beta subunit (EC 1.2.1.2)
Formate dehydrogenase O gamma subunit (EC 1.2.1.2)
Formate dehydrogenase-O, iron-sulfur subunit (EC 1.2.1.2)
Formate dehydrogenase-O, major subunit (EC 1.2.1.2)
Formate efflux transporter (TC 2.A.44 family)
Formate hydrogenlyase regulatory protein HycA
Formate hydrogenlyase subunit 3
Formate hydrogenlyase subunit 3/Multisubunit Na+/H+ antiporter, MnhD subunit
Formate hydrogenlyase subunit 4
Formate hydrogenlyase subunit 5
Formate hydrogenlyase subunit 7
Formate hydrogenlyase transcriptional activator
Formate--tetrahydrofolate ligase (EC 6.3.4.3)
Formiminoglutamic iminohydrolase (EC 3.5.3.13)
Forms the bulk of type IV secretion complex that spans outer membrane and periplasm (VirB9)
Formyltetrahydrofolate deformylase (EC 3.5.1.10)
Fructokinase (EC 2.7.1.4)
Fructose-1,6-bisphosphatase, GlpX type (EC 3.1.3.11)
Fructose-1,6-bisphosphatase, type I (EC 3.1.3.11)
Fructose-bisphosphate aldolase class I (EC 4.1.2.13)
Fructose-bisphosphate aldolase class II (EC 4.1.2.13)
Fructose-specific phosphocarrier protein HPr
Fructuronate transporter GntP
Fumarate and nitrate reduction regulatory protein
Fumarate hydratase class I (EC 4.2.1.2)
Fumarate hydratase class I, aerobic (EC 4.2.1.2)
Fumarate hydratase class I, anaerobic (EC 4.2.1.2)
Fumarate hydratase class II (EC 4.2.1.2)
Fumarate reductase subunit D

Glycerol-3-phosphate acyltransferase (EC 2.3.1.15)
Glycerol-3-phosphate dehydrogenase [NAD(P)+] (EC 1.1.1.94)
Glycerol-3-phosphate regulon repressor, DeoR family
Glycerol-3-phosphate transporter
Glycerophosphoryl diester phosphodiesterase (EC 3.1.4.46)
Glycine betaine transporter OpuD
Glycine cleavage system H protein
Glycine cleavage system transcriptional activator GcvA
Glycine cleavage system transcriptional antiactivator GcvR
Glycine dehydrogenase [decarboxylating] (glycine cleavage system P protein) (EC 1.4.4.2)
Glycine dehydrogenase [decarboxylating] (glycine cleavage system P2 protein) (EC 1.4.4.2)
Glycine oxidase ThiO (EC 1.4.3.19)
Glycine/D-amino acid oxidase (deaminating) in putrescine utilization cluster
Glycogen debranching enzyme (EC 3.2.1.-)
Glycogen phosphorylase (EC 2.4.1.1)
Glycogen synthase, ADP-glucose transglucosylase (EC 2.4.1.21)
Glycolate dehydrogenase (EC 1.1.99.14), FAD-binding subunit GlcE
Glycolate dehydrogenase (EC 1.1.99.14), iron-sulfur subunit GlcF
Glycolate dehydrogenase (EC 1.1.99.14), subunit GlcD
Glycolate oxidase (EC 1.1.3.15)
Glycosyl transferase, family 2
Glycyl-tRNA synthetase alpha chain (EC 6.1.1.14)
Glycyl-tRNA synthetase beta chain (EC 6.1.1.14)
Glyoxylate carboligase (EC 4.1.1.47)
Glyoxylate reductase (EC 1.1.1.26)
GMP reductase (EC 1.7.1.7)
GMP synthase [glutamine-hydrolyzing] (EC 6.3.5.2)
GTP cyclohydrolase I (EC 3.5.4.16) type 1

SSU ribosomal protein S4p (S9e)
SSU ribosomal protein S4p (S9e), mitochondrial
SSU ribosomal protein S5e (S7p)
SSU ribosomal protein S5p (S2e)
SSU ribosomal protein S5p (S2e), chloroplast
SSU ribosomal protein S5p (S2e), mitochondrial
SSU ribosomal protein S6e
SSU ribosomal protein S6p, mitochondrial
SSU ribosomal protein S7p (S5e), mitochondrial
SSU ribosomal protein S8e
SSU ribosomal protein S8p (S15Ae), chloroplast
SSU ribosomal protein S8p (S15Ae), mitochondrial
SSU ribosomal protein S9e (S4p)
SSU ribosomal protein S9p (S16e)
SSU ribosomal protein S9p (S16e), chloroplast
SSU ribosomal protein S9p (S16e), mitochondrial
SSU ribosomal protein VAR1, mitochondrial
Stage II sporulation protein B
Stage II sporulation protein D (SpoIID)
Stage II sporulation serine phosphatase for sigma-F activation (SpoIIE)
Stage III sporulation protein AA
Stage IV sporulation pro-sigma-K processing enzyme (SpoIVFB)
Stage V sporulation protein B
Stage VI sporulation protein D
StbD replicon stabilization protein (antitoxin to StbE)
Sterol regulatory element binding transcription factor 1
Streptococcal cell surface hemoprotein receptor Shr
Streptolysin S export transmembrane permease (SagI)

Fumarylacetoacetase (EC 3.7.1.2)
Fumarylacetoacetate (FAA) hydrolase (EC 4.1.1.68)
Fumarylacetoacetate hydrolase family protein
Functional role page for Anaerobic nitric oxide reductase transcription regulator NorR
G:T/U mismatch-specific uracil/thymine DNA-glycosylase
Galactitol-1-phosphate 5-dehydrogenase (EC 1.1.1.251)
Galactokinase (EC 2.7.1.6)
Galactonate dehydratase (EC 4.2.1.6)
Galactosamine-6-phosphate isomerase (EC 5.3.1.-)
Galactosamine-6-phosphate isomerase (galactosamine-6-phosphate deaminase) (EC 5.3.1.-)
Galactose operon repressor, GalR-LacI family of transcriptional regulators
Galactose-1-phosphate uridylyltransferase (EC 2.7.7.10)
Galactose/methyl galactoside ABC transport system, ATP-binding protein MglA (EC 3.6.3.17)
Galactose/methyl galactoside ABC transport system, D-galactose-binding periplasmic protein MglB (TC 3.A.1.2.3)
Galactose/methyl galactoside ABC transport system, permease protein MglC (TC 3.A.1.2.3)
Gallate dioxygenase (EC 1.13.11.57)
Gamma-aminobutyrate:alpha-ketoglutarate aminotransferase (EC 2.6.1.19)
Gamma-D-Glutamyl-meso-Diaminopimelate Amidase
Gamma-glutamyl phosphate reductase (EC 1.2.1.41)
Gamma-glutamyl-aminobutyraldehyde dehydrogenase (EC 1.2.1.-)
Gamma-glutamyl-putrescine oxidase (EC1.4.3.-)
Gamma-glutamyl-putrescine synthetase (EC 6.3.1.11)
Gamma-glutamyltranspeptidase (EC 2.3.2.2)
GDP-mannose 4,6-dehydratase (EC 4.2.1.47)
GDP-mannose pyrophosphatase YffH
General secretion pathway protein C
General secretion pathway protein F
General secretion pathway protein I

GTP cyclohydrolase I (EC 3.5.4.16) type 2
GTP cyclohydrolase II (EC 3.5.4.25)
GTP pyrophosphokinase (EC 2.7.6.5), (p)ppGpp synthetase I
GTP pyrophosphokinase (EC 2.7.6.5), (p)ppGpp synthetase II
GTP-binding and nucleic acid-binding protein YchF
GTP-binding protein EngA
GTP-binding protein EngB
GTP-binding protein Era
GTP-binding protein HflX
GTP-binding protein TypA/BipA
GTP-binding protein YqeH, required for biogenesis of 30S ribosome subunit
GTPase and tRNA-U34 5-formylation enzyme TrmE
Guanine deaminase (EC 3.5.4.3)
Guanine-hypoxanthine permease
Guanosine-3',5'-bis(diphosphate) 3'-pyrophosphohydrolase (EC 3.1.7.2)
Guanylate kinase (EC 2.7.4.8)
Hca operon (3-phenylpropionic acid catabolism) transcriptional activator HcaR
Heat shock protein 60 family chaperone GroEL
Heat shock protein 60 family co-chaperone GroES
Heat shock protein GrpE
Heat-inducible transcription repressor HrcA
Heavy metal RND efflux outer membrane protein, CzcC family
Heavy metal sensor histidine kinase
Helicase PriA essential for oriC/DnaA-independent DNA replication
Heme A synthase, cytochrome oxidase biogenesis protein Cox15-CtaA
Heme ABC transporter, ATPase component HmuV
Heme ABC transporter, cell surface heme and hemoprotein receptor HmuT

Streptothricin acetyltransferase, Streptomyces lavendulae type
Substrate-specific component NikM of nickel ECF transporter
Substrate-specific component RibU of riboflavin ECF transporter
Succinate dehydrogenase cytochrome b558 subunit
Succinate dehydrogenase cytochrome b560 subunit
Succinate dehydrogenase hydrophobic membrane anchor protein
Sulfhydrogenase II subunit a
Sulfhydrogenase II subunit b
Sulfite reductase, dissimilatory-type gamma subunit (EC 1.8.99.3)
Sulfite reduction-associated complex DsrMKJOP iron-sulfur protein DsrO (=HmeA)
Sulfite reduction-associated complex DsrMKJOP multiheme protein DsrJ (=HmeF)
Sulfite reduction-associated complex DsrMKJOP protein DsrM (= HmeC)
Sulfite reduction-associated complex DsrMKJOP protein DsrP (= HmeB)
sulfonate monooxygenase
Sulfur carrier protein adenylyltransferase ThiF
Suppression of copper sensitivity: putative copper binding protein ScsA
TcuA: flavoprotein used to oxidize tricarballylate to cis-aconitate
TcuC: integral membrane protein used to transport tricarballylate across the cell membrane
TcuR: regulates tcuABC genes used in utilization of tricarballylate
Tetrachloroethene reductive dehalogenase PceA membrane-bound subunit
Tetrachloroethene reductive dehalogenase TceA
Thiol:disulfide oxidoreductase TlpA
Thioredoxin reductase (EC 1.8.1.9)
Threonine catabolic operon transcriptional activator TdcR
Threonine dehydratase, catabolic (EC 4.3.1.19)
threonine dehydrogenase
tmRNA-binding protein SmpB

General secretion pathway protein L
Gentisate 1,2-dioxygenase (EC 1.13.11.4)
Geranyl-CoA carboxylase biotin-containing subunit (EC 6.4.1.5)
Geranylgeranyl diphosphate reductase (EC 1.3.1.83)
Geranylgeranyl pyrophosphate synthetase (EC 2.5.1.29)
Glucans biosynthesis glucosyltransferase H (EC 2.4.1.-)
Glucarate dehydratase (EC 4.2.1.40)
Glucitol operon repressor
Glucoamylase (EC 3.2.1.3)
Glucokinase (EC 2.7.1.2)
Gluconate dehydratase (EC 4.2.1.39)
Gluconate permease
Gluconate permease, Bsu4004 homolog
Gluconate transporter family protein
Gluconate utilization system Gnt-I transcriptional repressor
Gluconolactonase (EC 3.1.1.17)
Glucosamine--fructose-6-phosphate aminotransferase [isomerizing] (EC 2.6.1.16)
Glucosamine-6-phosphate deaminase (EC 3.5.99.6)
Glucose 1-dehydrogenase (EC 1.1.1.47)
Glucose dehydrogenase, PQQ-dependent (EC 1.1.5.2)
Glucose-1-phosphate adenylyltransferase (EC 2.7.7.27)
Glucose-1-phosphate cytidylyltransferase (EC 2.7.7.33)
Glucose-1-phosphate thymidylyltransferase (EC 2.7.7.24)
Glucose-6-phosphate 1-dehydrogenase (EC 1.1.1.49)
Glucose-6-phosphate isomerase (EC 5.3.1.9)
Glucose-methanol-choline (GMC) oxidoreductase:NAD binding site
Glucuronide transport facilitator UidC

Heme O synthase, protoheme IX farnesyltransferase (EC 2.5.1.-) COX10-CtaB
Heme-regulated cyclic AMP phosphodiesterase (EC 3.1.4.-)
Hemoglobin-like protein HbO
Hexuronate transporter
HflK protein
High-affinity branched-chain amino acid transport system permease protein LivH (TC 3.A.1.4.1)
High-affinity choline uptake protein BetT
High-affinity leucine-specific transport system, periplasmic binding protein LivK (TC 3.A.1.4.1)
Histidine ABC transporter, ATP-binding protein HisP (TC 3.A.1.3.1)
Histidine ABC transporter, permease protein HisM (TC 3.A.1.3.1)
Histidine ABC transporter, permease protein HisQ (TC 3.A.1.3.1)
Histidine ammonia-lyase (EC 4.3.1.3)
Histidine utilization repressor
Histidinol dehydrogenase (EC 1.1.1.23)
Histidinol-phosphatase (EC 3.1.3.15)
Histidinol-phosphate aminotransferase (EC 2.6.1.9)
HmrA protein involved in methicillin resistance
Holliday junction DNA helicase RuvA
Holliday junction DNA helicase RuvB
Holo-[acyl-carrier protein] synthase (EC 2.7.8.7)
Homocitrate synthase (EC 2.3.3.14)
Homocysteine S-methyltransferase (EC 2.1.1.10)
Homogentisate 1,2-dioxygenase (EC 1.13.11.5)
Homolog of fucose/glucose/galactose permeases
Homoprotocatechuate degradative operon repressor
Homoserine dehydrogenase (EC 1.1.1.3)
Homoserine kinase (EC 2.7.1.39)

TolA protein
tolB protein precursor, periplasmic protein involved in the tonb-independent uptake of group A colicins
Toluate 1,2-dioxygenase alpha subunit
Toluene-4-monooxygenase, subunit TmoD
TonB system biopolymer transport component
TonB-dependent hemin , ferrichrome receptor
Topoisomerase IV subunit A (EC 5.99.1.-)
Topoisomerase IV subunit B (EC 5.99.1.-)
Toxin co-regulated pilus biosynthesis protein B
Toxin co-regulated pilus biosynthesis protein C, outer membrane protein
Toxin co-regulated pilus biosynthesis protein D
Toxin co-regulated pilus biosynthesis protein F, putative outer membrane channel for TcpA extrusion
Toxin co-regulated pilus biosynthesis protein H, transcriptional activator of ToxT promoter
Toxin co-regulated pilus biosynthesis protein I, chemoreceptor, negative regulator of TcpA
Toxin co-regulated pilus biosynthesis protein Q
Toxin co-regulated pilus biosynthesis protein R
Toxin co-regulated pilus biosynthesis protein S
TPR repeat containing exported protein
Trans-2,cis-3-Decenoyl-ACP isomerase
Transaldolase (EC 2.2.1.2)
Transcription termination protein NusA
Transcriptional regulator BkdR of isoleucine and valine catabolism operon
Transcriptional regulator CsgD for 2nd curli operon
Transcriptional regulator GabR of GABA utilization (GntR family with aminotransferase-like domain)
Transcriptional regulator LasR
Transcriptional regulator NfxB
Transcriptional regulator of mannoside utilization, variant 2, LacI family

Glucuronide transporter UidB
Glutamate 5-kinase (EC 2.7.2.11)
Glutamate N-acetyltransferase (EC 2.3.1.35) / N-acetylglutamate synthase (EC 2.3.1.1)
Glutamate racemase (EC 5.1.1.3)
Glutamate synthase [NADPH] large chain (EC 1.4.1.13)
Glutamate synthase [NADPH] small chain (EC 1.4.1.13)
Glutamate--cysteine ligase (EC 6.3.2.2)
Glutamate--cysteine ligase (EC 6.3.2.2), divergent, of Alpha- and Beta-proteobacteria type
Glutamate-1-semialdehyde aminotransferase (EC 5.4.3.8)
Glutamate-ammonia-ligase adenylyltransferase (EC 2.7.7.42)
Glutaminase (EC 3.5.1.2)
glutamine synthetase family protein
Glutamine synthetase type I (EC 6.3.1.2)
Glutamine synthetase type III, GlnN (EC 6.3.1.2)
Glutamine-dependent 2-keto-4-methylthiobutyrate transaminase
Glutaminyl-tRNA synthetase (EC 6.1.1.18)
glutamyl-Q-tRNA synthetase
Glutamyl-tRNA reductase (EC 1.2.1.70)
Glutamyl-tRNA synthetase (EC 6.1.1.17)
Glutamyl-tRNA(Gln) amidotransferase subunit A (EC 6.3.5.7)
Glutamyl-tRNA(Gln) amidotransferase subunit B (EC 6.3.5.7)
Glutamyl-tRNA(Gln) amidotransferase subunit C (EC 6.3.5.7)
Glutamyl-tRNA(Gln) synthetase (EC 6.1.1.24)
Glutaredoxin 1
Glutaredoxin 2
Glutaryl-CoA dehydrogenase (EC 1.3.99.7)
Glutathione peroxidase (EC 1.11.1.9)

Homoserine O-acetyltransferase (EC 2.3.1.31)
HoxN/HupN/NixA family nickel/cobalt transporter
HTH domain protein SA1665, binds to mecA promoter region
HTH-type transcriptional regulator appY
HTH-type transcriptional regulator BetI
HTH-type transcriptional regulator IlvY
HtrA protease/chaperone protein
HupE-UreJ family metal transporter
Hydantoin permease
Hydantoin racemase (EC 5.1.99.-)
Hydantoinase
Hydrogen peroxide-inducible genes activator
Hydrolase in polyol utilization gene cluster, haloacid dehalogenase-like family
Hydrolase, alpha/beta fold family functionally coupled to Phosphoribulokinase
Hydroxyacylglutathione hydrolase (EC 3.1.2.6)
Hydroxyethylthiazole kinase (EC 2.7.1.50)
Hydroxymethylglutaryl-CoA lyase (EC 4.1.3.4)
Hydroxymethylglutaryl-CoA synthase (EC 2.3.3.10)
Hydroxymethylpyrimidine ABC transporter, substrate-binding component
Hydroxymethylpyrimidine ABC transporter, transmembrane component
Hydroxymethylpyrimidine phosphate synthase ThiC (EC 4.1.99.17)
Hydroxypyruvate reductase (EC 1.1.1.81)
Hypothetical lactonase in carbon monoxide dehydrogenase cluster
Hypothetical nudix hydrolase YeaB
Hypothetical protein GlcG in glycolate utilization operon
Hypothetical protein i Rubrerythrin cluster
Hypothetical protein YagQ
Hypothetical radical SAM family enzyme in heat shock gene cluster, similarity

Transcriptional regulator of rhamnose utilization, GntR family
Transcriptional regulator of rhamnose utilization, LysR family
Transcriptional regulator RhlR
Transcriptional regulator RutR of pyrimidine catabolism (TetR family)
Transcriptional regulator, MerR family, near polyamine transporter
Transcriptional regulator, TetR family, associated with agmatine catabolism
Transcriptional repressor of aerobactin receptor iutR
Transcriptional repressor of the lac operon
Transketolase (EC 2.2.1.1)
Transketolase, N-terminal section (EC 2.2.1.1)
Transmembrane component NikQ of energizing module of nickel ECF transporter
Transport ATP-binding protein CydC
Transport ATP-binding protein CydCD
Triphosphoribosyl-dephospho-CoA synthetase (EC 2.7.8.25)
tRNA 2-thiouridine synthesizing protein E (EC 2.8.1.-)
tRNA methylase Trm12p Wyeosine biosynthesis
tRNA methylase Yol141w homolog Wyeosine biosynthesis
tRNA-Asx
tRNA-His-GUG
TTE0858 replicon stabilization protein (antitoxin to TTE0859)
Twin-arginine translocation protein TatB
Twin-arginine translocation protein TatC
Twin-arginine translocation protein TatE
Two-component response regulator colocalized with HrtAB transporter
Two-component system histidine kinase
Two-component system response regulator
Type I secretion outer membrane protein, TolC precursor
Type IV secretion system protein VirD4

Glutathione reductase (EC 1.8.1.7)
Glutathione S-transferase (EC 2.5.1.18)
Glutathione S-transferase family protein
Glutathione S-transferase, omega (EC 2.5.1.18)
Glutathione S-transferase, unnamed subgroup (EC 2.5.1.18)
Glutathione S-transferase, unnamed subgroup 2 (EC 2.5.1.18)
Glutathione synthetase (EC 6.3.2.3)
Glutathione-dependent formaldehyde-activating enzyme (EC 4.4.1.22)
Glutathione-regulated potassium-efflux system ancillary protein KefF
Glutathione-regulated potassium-efflux system ATP-binding protein
Glutathione-regulated potassium-efflux system protein KefB
Glutathione-regulated potassium-efflux system protein KefC
Glutathionylspermidine amidohydrolase (EC 3.5.1.78)
Glyceraldehyde-3-phosphate ketol-isomerase (EC 5.3.1.1)
Glycerate kinase (EC 2.7.1.31)
Glycerol dehydrogenase (EC 1.1.1.6)
Glycerol kinase (EC 2.7.1.30)
Glycerol uptake facilitator protein
Glycerol-3-phosphate ABC transporter, ATP-binding protein UgpC (TC 3.A.1.1.3)
Glycerol-3-phosphate ABC transporter, periplasmic glycerol-3-phosphate-binding protein (TC 3.A.1.1.3)
Glycerol-3-phosphate ABC transporter, permease protein UgpA (TC 3.A.1.1.3)
Glycerol-3-phosphate ABC transporter, permease protein UgpE (TC 3.A.1.1.3)
Glycerol-3-phosphate acyltransferase (EC 2.3.1.15)
Glycerol-3-phosphate dehydrogenase (EC 1.1.5.3)
Glycerol-3-phosphate dehydrogenase [NAD(P)+] (EC 1.1.1.94)
Glycerol-3-phosphate regulon repressor, DeoR family
Glycerophosphoryl diester phosphodiesterase (EC 3.1.4.46)
Glycine cleavage system H protein

with CPO of BS HemN-type
Hypothetical radical SAM family enzyme, NOT coproporphyrinogen III oxidase, oxygen-independent
Hypothetical YciO protein, TsaC/YrdC paralog
Hypoxanthine-guanine phosphoribosyltransferase (EC 2.4.2.8)
Imidazole glycerol phosphate synthase amidotransferase subunit (EC 2.4.2.-)
Imidazole glycerol phosphate synthase cyclase subunit (EC 4.1.3.-)
Imidazoleglycerol-phosphate dehydratase (EC 4.2.1.19)
Imidazolonepropionase (EC 3.5.2.7)
IMP cyclohydrolase (EC 3.5.4.10)
IncF plasmid conjugative transfer DNA-nicking and unwinding protein TraI
IncF plasmid conjugative transfer mating signal transduction protein TraM
IncF plasmid conjugative transfer pilus assembly protein TraC
IncF plasmid conjugative transfer pilus assembly protein TraE
IncF plasmid conjugative transfer pilus assembly protein TraF
IncF plasmid conjugative transfer pilus assembly protein TraH
IncF plasmid conjugative transfer pilus assembly protein TraK
IncF plasmid conjugative transfer pilus assembly protein TraU
IncF plasmid conjugative transfer pilus assembly protein TraW
IncF plasmid conjugative transfer protein TrbC
IncF plasmid conjugative transfer protein TrbI
IncF plasmid conjugative transfer regulator TraY
IncW plasmid conjugative relaxase protein TrwC (TraI homolog)
Indole-3-glycerol phosphate synthase (EC 4.1.1.48)
Indoleacetamide hydrolase (EC 3.5.1.-)
Indoleamine 2,3-dioxygenase (EC 1.13.11.52)
Indolepyruvate ferredoxin oxidoreductase, alpha and beta subunits
Indolepyruvate oxidoreductase subunit IorA (EC 1.2.7.8)

Type IV secretion/competence protein (VirB10)
Type IV secretion/competence protein (VirB9)
Ubiquinone/menaquinone biosynthesis methyltransferase UbiE (EC 2.1.1.-)
UDP-6-deoxy-AltdiNAc hydrolase (PseG, third step of pseudaminic acid biosynthesis)
UDP-N-acetylglucosamine 1-carboxyvinyltransferase (EC 2.5.1.7)
UDP-N-acetylmuramoylalanine--D-glutamate ligase (EC 6.3.2.9)
UDP-N-acetylmuramoylalanyl-D-glutamyl-2,6-diaminopimelate--D-alanyl-D-alanine ligase (EC 6.3.2.10)
Uncharacterized iron compound ABC uptake transporter, ATP-binding protein
Uncharacterized iron compound ABC uptake transporter, permease protein
Uncharacterized iron compound ABC uptake transporter, substrate-binding protein
Uncharacterized monothiol glutaredoxin ycf64-like
Uncharacterized protein conserved in archaea (DUF531)
Uncharacterized protein YkvT, NOT involved in spore germination
Uncharacterized spore germination protein YndD
Uncharacterized spore germination protein YndE
Uncharacterized spore germination protein YndF
Universal stress protein E
Universal stress protein F
Universal stress protein family
Universal stress protein family 2
Universal stress protein family 6
Ureidoglycolate/malate/sulfolactate dehydrogenase family (EC 1.1.1.-)
Uroporphyrinogen-III synthase (EC 4.2.1.75)
Vancomycin response regulator VanR
VapC7 antibacterial toxin protein
Various polyols ABC transporter, periplasmic substrate-binding protein

Glycine cleavage system transcriptional activator
Glycine cleavage system transcriptional activator GcvA
Glycine cleavage system transcriptional antiactivator GcvR
Glycine dehydrogenase [decarboxylating] (glycine cleavage system P protein) (EC 1.4.4.2)
Glycine dehydrogenase [decarboxylating] (glycine cleavage system P1 protein) (EC 1.4.4.2)
Glycine dehydrogenase [decarboxylating] (glycine cleavage system P2 protein) (EC 1.4.4.2)
Glycine oxidase ThiO (EC 1.4.3.19)
Glycine/D-amino acid oxidase (deaminating) in putrescine utilization cluster
Glycogen debranching enzyme (EC 3.2.1.-)
Glycogen phosphorylase (EC 2.4.1.1)
Glycogen synthase, ADP-glucose transglucosylase (EC 2.4.1.21)
Glycolate dehydrogenase (EC 1.1.99.14), FAD-binding subunit GlcE
Glycolate dehydrogenase (EC 1.1.99.14), iron-sulfur subunit GlcF
Glycolate dehydrogenase (EC 1.1.99.14), subunit GlcD
Glycolate oxidase (EC 1.1.3.15)
Glycolate permease
Glycolate utilization operon transcriptional activator GlcC
Glycosyl hydrolase YegX, family 25
Glycosyl transferase, family 2
Glycosyltransferase LafB, responsible for the formation of Gal-Glc-DAG
Glycyl-tRNA synthetase alpha chain (EC 6.1.1.14)
Glycyl-tRNA synthetase beta chain (EC 6.1.1.14)
Glyoxylate carboligase (EC 4.1.1.47)
Glyoxylate reductase (EC 1.1.1.26)
GMP reductase (EC 1.7.1.7)
GMP synthase [glutamine-hydrolyzing] (EC 6.3.5.2)

Inner membrane component of tripartite multidrug resistance system

Inner membrane protein CreD

Inner membrane protein CreD-like protein

Inner membrane protein forms channel for type IV secretion of T-DNA complex (VirB10)

Inner membrane protein forms channel for type IV secretion of T-DNA complex (VirB8)

Inner membrane protein translocase component YidC, long form

Inorganic pyrophosphatase (EC 3.6.1.1)

Inosine-5'-monophosphate dehydrogenase (EC 1.1.1.205)

Inosine-guanosine kinase (EC 2.7.1.73)

Inosine-uridine preferring nucleoside hydrolase (EC 3.2.2.1)

Inositol-1-monophosphatase (EC 3.1.3.25)

Inosose dehydratase (EC 4.2.1.44)

Inosose isomerase (EC 5.3.99.-)

Integral inner membrane protein of type IV secretion complex (VirB6)

Integral membrane protein YggT, involved in response to extracytoplasmic stress (osmotic shock)

Integration host factor alpha subunit

Integration host factor beta subunit

Intracellular PHB depolymerase (EC 3.1.1.-)

Intradiol ring-cleavage dioxygenase (EC 1.13.11.1)

iron-chelator utilization protein

Iron-responsive regulator Irr

Iron-sulfur cluster assembly protein SufB

Iron-sulfur cluster regulator IscR

Isoaspartyl dipeptidase (EC 3.4.19.5)

Isocitrate dehydrogenase [NADP] (EC 1.1.1.42)

Isocitrate lyase (EC 4.1.3.1)

Isoquinoline 1-oxidoreductase alpha subunit (EC 1.3.99.16)

Isoquinoline 1-oxidoreductase beta subunit (EC 1.3.99.16)

Ketol-acid reductoisomerase (EC 1.1.1.86)

Various polyols ABC transporter, permease component 1

Various polyols ABC transporter, permease component 2

virulence cluster protein B VclB

Vulnibactin utilization protein VuuB

Xanthosine permease

Xanthosine permease, C-terminus

Xylose ABC transporter, periplasmic xylose-binding protein XylF

YhfA Protein in tra region of some IncF plasmids

Ync

Ynd

Zinc metalloproteinase precursor (EC 3.4.24.29)

Zinc resistance-associated protein

Zinc transporter ZitB

GTP cyclohydrolase I (EC 3.5.4.16) type 1

GTP cyclohydrolase II (EC 3.5.4.25)

GTP pyrophosphokinase (EC 2.7.6.5), (p)ppGpp synthetase I

GTP pyrophosphokinase (EC 2.7.6.5), (p)ppGpp synthetase I

GTP pyrophosphokinase (EC 2.7.6.5), (p)ppGpp synthetase II

GTP-binding and nucleic acid-binding protein YchF

GTP-binding protein EngA

GTP-binding protein EngB

GTP-binding protein Era

GTP-binding protein HflX

GTP-binding protein related to HflX

GTP-binding protein TypA/BipA

GTPase and tRNA-U34 5-formylation enzyme TrmE

Guanine deaminase (EC 3.5.4.3)

Guanosine-3',5'-bis(diphosphate) 3'-pyrophosphohydrolase (EC 3.1.7.2)

Guanylate kinase (EC 2.7.4.8)

Hca operon (3-phenylpropionic acid catabolism) transcriptional activator HcaR

Heat shock protein 60 family chaperone GroEL

Heat shock protein 60 family co-chaperone GroES

Heat shock protein GrpE

Heat-inducible transcription repressor HrcA

Heavy metal resistance transcriptional regulator HmrR

Heavy metal RND efflux outer membrane protein, CzcC family

Helicase PriA essential for oriC/DnaA-independent DNA replication

Heme A synthase, cytochrome oxidase biogenesis protein Cox15-CtaA

Heme ABC transporter, cell surface heme and hemoprotein receptor HmuT

Heme ABC transporter, permease protein HmuU

Heme biosynthesis protein related to NirL and NirH

Heme O synthase, protoheme IX farnesyltransferase (EC 2.5.1.-) COX10-CtaB

Ku domain protein
Kup system potassium uptake protein
Kynureninase (EC 3.7.1.3)
Kynurenine formamidase, bacterial (EC 3.5.1.9)
L-alanine:glyoxylate aminotransferase (EC 2.6.1.44)
L-allo-threonine aldolase (EC 2.1.2.1)
L-arabinose isomerase (EC 5.3.1.4)
L-arabinose transport ATP-binding protein AraG (TC 3.A.1.2.2)
L-arabinose transport system permease protein (TC 3.A.1.2.2)
L-arabinose-binding periplasmic protein precursor AraF (TC 3.A.1.2.2)
L-arabonate dehydratase (EC 4.2.1.25)
L-asparaginase (EC 3.5.1.1)
L-aspartate oxidase (EC 1.4.3.16)
L-carnitine/gamma-butyrobetaine antiporter
L-fuconate dehydratase (EC 4.2.1.68)
L-fuculokinase (EC 2.7.1.51)
L-lactate dehydrogenase (EC 1.1.1.27)
L-lactate dehydrogenase (EC 1.1.2.3)
L-lactate permease
L-pipecolate dehydrogenase (EC 1.5.99.3)
L-proline glycine betaine ABC transport system permease protein ProV (TC 3.A.1.12.1)
L-proline glycine betaine ABC transport system permease protein ProW (TC 3.A.1.12.1)
L-proline glycine betaine binding ABC transporter protein ProX (TC 3.A.1.12.1)
L-Proline/Glycine betaine transporter ProP
L-rhamnose operon transcriptional activator RhaR
L-ribulose-5-phosphate 4-epimerase UlaF (EC 5.1.3.4) (L-ascorbate utilization protein F)
L-serine dehydratase (EC 4.3.1.17)
L-serine dehydratase, alpha subunit (EC 4.3.1.17)
L-seryl-tRNA(Sec) selenium transferase (EC 2.9.1.1)
L-threonine 3-dehydrogenase (EC 1.1.1.103)

Heme-regulated cyclic AMP phosphodiesterase (EC 3.1.4.-)
Hexose phosphate transport protein UhpT
Hexose phosphate uptake regulatory protein UhpC
Hexuronate transporter
HflC protein
HflK protein
High-affinity branched-chain amino acid transport system permease protein LivH (TC 3.A.1.4.1)
High-affinity choline uptake protein BetT
High-affinity gluconate transporter GntT
High-affinity leucine-specific transport system, periplasmic binding protein LivK (TC 3.A.1.4.1)
Histidine ABC transporter, ATP-binding protein HisP (TC 3.A.1.3.1)
Histidine ABC transporter, permease protein HisM (TC 3.A.1.3.1)
Histidine ammonia-lyase (EC 4.3.1.3)
Histidine transport protein (permease)
Histidinol dehydrogenase (EC 1.1.1.23)
Histidinol-phosphatase (EC 3.1.3.15)
Histidinol-phosphatase [alternative form] (EC 3.1.3.15)
Histidinol-phosphate aminotransferase (EC 2.6.1.9)
Hnr protein
Holliday junction DNA helicase RuvA
Holliday junction DNA helicase RuvB
Holo-[acyl-carrier protein] synthase (EC 2.7.8.7)
Homocitrate synthase (EC 2.3.3.14)
Homogentisate 1,2-dioxygenase (EC 1.13.11.5)
Homolog of fucose/glucose/galactose permeases
Homoprotocatechuate degradative operon repressor
Homoserine dehydrogenase (EC 1.1.1.3)
Homoserine kinase (EC 2.7.1.39)
Homoserine O-acetyltransferase (EC 2.3.1.31)
Homoserine O-succinyltransferase (EC 2.3.1.46)

L-threonine 3-O-phosphate decarboxylase (EC 4.1.1.81)
L,D-transpeptidase YcbB
L,D-transpeptidase YnhG
Lactate-responsive regulator LldR in Actinobacteria, GntR family
Lacto-N-biose phosphorylase (EC 2.4.1.211)
Lactose permease
Lactoylglutathione lyase (EC 4.4.1.5)
Large exoproteins involved in heme utilization or adhesion
Large-conductance mechanosensitive channel
Late competence protein ComEB
Leucine dehydrogenase (EC 1.4.1.9)
Leucine-responsive regulatory protein, regulator for leucine (or lrp) regulon and high-affinity branched-chain amino acid transport system
Leucyl/phenylalanyl-tRNA--protein transferase (EC 2.3.2.6)
Light-independent protochlorophyllide reductase subunit N (EC 1.18.-.-)
Lipid A biosynthesis lauroyl acyltransferase (EC 2.3.1.-)
Lipid A export ATP-binding/permease protein MsbA
Lipid-A-disaccharide synthase (EC 2.4.1.182)
Lipoate synthase
Lipocalin-related protein and Bos/Can/Equ allergen
Lipopolysaccharide ABC transporter, ATP-binding protein LptB
Lipopolysaccharide biosynthesis protein WzxC
Lipopolysaccharide core biosynthesis glycosyltransferase, group 2 family protein (EC 2.4.1.-)
Lipopolysaccharide heptosyltransferase I (EC 2.4.1.-)
Lipoprotein signal peptidase (EC 3.4.23.36)
Long-chain-fatty-acid--CoA ligase (EC 6.2.1.3)
Low-affinity inorganic phosphate transporter
Low-specificity L-threonine aldolase (EC 4.1.2.48)
LptA, protein essential for LPS transport across the periplasm
LrgA-associated membrane protein LrgB

HoxN/HupN/NixA family nickel/cobalt transporter
HPr kinase/phosphorylase (EC 2.7.1.-) (EC 2.7.4.-)
HspR, transcriptional repressor of DnaK operon
HTH-type transcriptional regulator IlvY
HtrA protease/chaperone protein
Hybrid sensory histidine kinase in two-component regulatory system with EvgA
Hydantoin permease
Hydantoin racemase (EC 5.1.99.-)
Hydrogen peroxide-inducible genes activator
Hydrogenase-4 component C
Hydrogenase-4 component E
Hydrolase, alpha/beta fold family functionally coupled to Phosphoribulokinase
Hydroxyacylglutathione hydrolase (EC 3.1.2.6)
Hydroxyethylthiazole kinase (EC 2.7.1.50)
Hydroxylamine reductase (EC 1.7.-.-)
Hydroxymethylglutaryl-CoA lyase (EC 4.1.3.4)
Hydroxymethylglutaryl-CoA reductase (EC 1.1.1.34)
Hydroxymethylpyrimidine ABC transporter, ATPase component
Hydroxymethylpyrimidine ABC transporter, substrate-binding component
Hydroxymethylpyrimidine ABC transporter, transmembrane component
Hydroxymethylpyrimidine phosphate synthase ThiC (EC 4.1.99.17)
Hydroxypyruvate reductase (EC 1.1.1.81)
Hypothetical glycoside hydrolase, family 43, similar to arabinosidase
Hypothetical lactonase in carbon monoxide dehydrogenase cluster
Hypothetical nudix hydrolase YeaB
Hypothetical protein GlcG in glycolate utilization operon
Hypothetical protein TdcF in cluster with anaerobic degradation of L-threonine to propionate
hypothetical protein that often co-occurs with aconitase
Hypothetical protein VC0266 (sugar utilization related?)

LsrR, transcriptional repressor of lsr operon	Hypothetical radical SAM family enzyme in heat shock gene cluster, similarity with CPO of BS HemN-type
LSU m3Psi1915 methyltransferase RlmH	Hypothetical transmembrane protein coupled to NADH-ubiquinone oxidoreductase chain 5 homolog
LSU m5C1962 methyltransferase RlmI	Hypothetical YciO protein, TsaC/YrdC paralog
LSU ribosomal protein L10p (P0)	Hypothetical YciO protein, TsaC/YrdC paralog
LSU ribosomal protein L11p (L12e)	Hypothetical, distant similarity with heme-degrading oxygenase IsdG
LSU ribosomal protein L13p (L13Ae)	Hypothetical, related to broad specificity phosphatases COG0406
LSU ribosomal protein L14p (L23e)	Hypoxanthine-guanine phosphoribosyltransferase (EC 2.4.2.8)
LSU ribosomal protein L15p (L27Ae)	Imidazole glycerol phosphate synthase amidotransferase subunit (EC 2.4.2.-)
LSU ribosomal protein L16p (L10e)	Imidazole glycerol phosphate synthase cyclase subunit (EC 4.1.3.-)
LSU ribosomal protein L17p	Imidazoleglycerol-phosphate dehydratase (EC 4.2.1.19)
LSU ribosomal protein L19p	Imidazolonepropionase (EC 3.5.2.7)
LSU ribosomal protein L1p (L10Ae)	IMP cyclohydrolase (EC 3.5.4.10)
LSU ribosomal protein L20p	IncF plasmid conjugative transfer DNA-nicking and unwinding protein TraI
LSU ribosomal protein L22p (L17e)	IncF plasmid conjugative transfer pilus assembly protein TraB
LSU ribosomal protein L24p (L26e)	IncF plasmid conjugative transfer pilus assembly protein TraC
LSU ribosomal protein L25p	IncF plasmid conjugative transfer pilus assembly protein TraH
LSU ribosomal protein L27p	IncF plasmid conjugative transfer pilus assembly protein TraL
LSU ribosomal protein L29p (L35e)	IncF plasmid conjugative transfer protein TraD
LSU ribosomal protein L2p (L8e)	IncF plasmid conjugative transfer protein TraG
LSU ribosomal protein L30p (L7e)	IncF plasmid conjugative transfer protein TraN
LSU ribosomal protein L31p	IncF plasmid conjugative transfer protein TrbC
LSU ribosomal protein L32p	IncF plasmid conjugative transfer regulator TraY
LSU ribosomal protein L33p, zinc-independent	IncQ plasmid conjugative transfer protein TraP
LSU ribosomal protein L35p	IncW plasmid conjugative protein TrwB (TraD homolog)
LSU ribosomal protein L3p (L3e)	IncW plasmid conjugative relaxase protein TrwC (TraI homolog)
LSU ribosomal protein L4p (L1e)	Indole-3-glycerol phosphate synthase (EC 4.1.1.48)
LSU ribosomal protein L5p (L11e)	Indoleacetamide hydrolase (EC 3.5.1.-)
LSU ribosomal protein L6p (L9e)	Indoleamine 2,3-dioxygenase (EC 1.13.11.52)
LSU ribosomal protein L7/L12 (P1/P2)	Indolepyruvate ferredoxin oxidoreductase, alpha and beta subunits
LSU ribosomal protein L9p	Indolepyruvate oxidoreductase subunit IorA (EC 1.2.7.8)

Lysine 2,3-aminomutase (EC 5.4.3.2)
Lysine decarboxylase (EC 4.1.1.18)
Lysine decarboxylase, inducible (EC 4.1.1.18)
Lysine-arginine-ornithine-binding periplasmic protein precursor (TC 3.A.1.3.1)
Lysine/cadaverine antiporter membrane protein CadB
Lysophospholipase (EC 3.1.1.5)
Lysophospholipase L2 (EC 3.1.1.5)
LysR family transcriptional regulator near succinyl-CoA:3-ketoacid-coenzyme A transferase
LysR family transcriptional regulator PA0133
LysR family transcriptional regulator PA3398
LysR family transcriptional regulator YbhD
LysR family transcriptional regulator YfeR
LysR family transcriptional regulator YfiE
LysR family transcriptional regulator YneJ
LysR family transcriptional regulator YnfL
Lysyl-tRNA synthetase (class I) (EC 6.1.1.6)
Lysyl-tRNA synthetase (class II) (EC 6.1.1.6)
Lysyl-tRNA synthetase (class II) related protein found fused to membrane protein
LytH protein involved in methicillin resistance
Macrolide export ATP-binding/permease protein MacB (EC 3.6.3.-)
Magnesium and cobalt efflux protein CorC
Magnesium and cobalt transport protein CorA
Major myo-inositol transporter IolT
Malate dehydrogenase (EC 1.1.1.37)
Malate Na(+) symporter
Malate synthase (EC 2.3.3.9)
Malate synthase G (EC 2.3.3.9)
Malate:quinone oxidoreductase (EC 1.1.5.4)
Maleylacetoacetate isomerase (EC 5.2.1.2)
Malonate decarboxylase alpha subunit

Indolepyruvate oxidoreductase subunit IorB (EC 1.2.7.8)
Inner membrane ABC transporter permease protein YcjO
Inner membrane ABC transporter permease protein YcjP
Inner membrane component of tripartite multidrug resistance system
Inner membrane protein forms channel for type IV secretion of T-DNA complex (VirB10)
Inner membrane protein forms channel for type IV secretion of T-DNA complex (VirB3)
Inner membrane protein forms channel for type IV secretion of T-DNA complex (VirB8)
Inner membrane protein translocase component YidC, long form
Inner membrane protein YphA
Inner membrane transport protein YeiM, in cluster with pseudouridine metabolism operon
Inorganic pyrophosphatase (EC 3.6.1.1)
Inosine-5'-monophosphate dehydrogenase (EC 1.1.1.205)
Inosine-guanosine kinase (EC 2.7.1.73)
Inosine-uridine preferring nucleoside hydrolase (EC 3.2.2.1)
Inositol transport system permease protein
Inositol-1-monophosphatase (EC 3.1.3.25)
Inosose dehydratase (EC 4.2.1.44)
Inosose isomerase (EC 5.3.99.-)
Integral inner membrane protein of type IV secretion complex (VirB6)
Integral membrane protein YggT, involved in response to extracytoplasmic stress (osmotic shock)
Integration host factor alpha subunit
Integration host factor beta subunit
Integron integrase IntIPac
intein-containing
internalin, putative
Intracellular PHB depolymerase (EC 3.1.1.-)
Intradiol ring-cleavage dioxygenase (EC 1.13.11.1)
Iron binding protein IscA for iron-sulfur cluster assembly
iron-chelator utilization protein
Iron-responsive regulator Irr

Malonyl CoA-acyl carrier protein transacylase (EC 2.3.1.39)
Malonyl-CoA decarboxylase (EC 4.1.1.9)
Malto-oligosyltrehalose synthase (EC 5.4.99.15)
Malto-oligosyltrehalose trehalohydrolase (EC 3.2.1.141)
Maltoporin (maltose/maltodextrin high-affinity receptor, phage lambda receptor protein)
Maltose operon transcriptional repressor MalR, LacI family
Maltose phosphorylase (EC 2.4.1.8)
Maltose regulon regulatory protein MalI (repressor for malXY)
Maltose/maltodextrin ABC transporter, permease protein MalF
Maltose/maltodextrin ABC transporter, permease protein MalG
Maltose/maltodextrin ABC transporter, substrate binding periplasmic protein MalE
Maltose/maltodextrin transport ATP-binding protein MalK (EC 3.6.3.19)
Malyl-CoA lyase (EC 4.1.3.24)
Manganese transport protein MntH
Manganese-dependent inorganic pyrophosphatase (EC 3.6.1.1)
Mannonate dehydratase (EC 4.2.1.8)
Mannose-1-phosphate guanylyltransferase (GDP) (EC 2.7.7.22)
Mannose-6-phosphate isomerase (EC 5.3.1.8)
Melibiose operon regulatory protein
membrane c-type cytochrome cy
Membrane fusion component of tripartite multidrug resistance system
Membrane fusion protein of RND family multidrug efflux pump
Membrane-associated zinc metalloprotease
Membrane-bound lytic murein transglycosylase B precursor (EC 3.2.1.-)
Membrane-bound lytic murein transglycosylase D precursor (EC 3.2.1.-)
Membrane-bound lytic murein transglycosylase E (EC 3.2.1.-)
Menaquinone via futalosine step 3
Menaquinone-specific isochorismate synthase (EC 5.4.4.2)
Mercuric resistance operon regulatory protein
Iron-sulfur cluster assembly protein SufB
Iron-sulfur cluster regulator IscR
Iron(III) dicitrate transport system permease protein FecC (TC 3.A.1.14.1)
Isoaspartyl dipeptidase (EC 3.4.19.5)
Isochorismatase (EC 3.3.2.1) of siderophore biosynthesis
Isochorismate synthase (EC 5.4.4.2) [enterobactin] siderophore
Isocitrate dehydrogenase [NADP] (EC 1.1.1.42)
Isocitrate lyase (EC 4.1.3.1)
Isopentenyl-diphosphate delta-isomerase (EC 5.3.3.2)
Isopentenyl-diphosphate delta-isomerase, FMN-dependent (EC 5.3.3.2)
Isoquinoline 1-oxidoreductase alpha subunit (EC 1.3.99.16)
Isoquinoline 1-oxidoreductase beta subunit (EC 1.3.99.16)
Kef-type transport system 2 (probable substrate potassium), subunit 2
Ketoglutarate semialdehyde dehydrogenase (EC 1.2.1.26)
Ketol-acid reductoisomerase (EC 1.1.1.86)
Ku domain protein
Kup system potassium uptake protein
Kynureninase (EC 3.7.1.3)
Kynurenine formamidase, bacterial (EC 3.5.1.9)
L-2-keto-3-deoxyarabonate dehydratase (EC 4.2.1.43)
L-alanine-DL-glutamate epimerase
L-arabinolactonase (EC 3.1.1.15)
L-arabinose isomerase (EC 5.3.1.4)
L-arabinose transport ATP-binding protein AraG (TC 3.A.1.2.2)
L-arabinose transport system permease protein (TC 3.A.1.2.2)
L-arabinose-binding periplasmic protein precursor AraF (TC 3.A.1.2.2)
L-arabonate dehydratase (EC 4.2.1.25)
L-asparaginase (EC 3.5.1.1)
L-aspartate oxidase (EC 1.4.3.16)

Metal-dependent hydrolase involved in phosphonate metabolism
Metal-dependent hydrolase YbeY, involved in rRNA and/or ribosome maturation and assembly
Metal-dependent hydrolases of the beta-lactamase superfamily I
Methane monooxygenase component A beta chain (EC 1.14.13.25)
Methanol dehydrogenase large subunit protein (EC 1.1.99.8)
Methenyltetrahydrofolate cyclohydrolase (EC 3.5.4.9)
Methionine ABC transporter ATP-binding protein
Methionine ABC transporter permease protein
Methionine ABC transporter substrate-binding protein
Methionine aminotransferase, PLP-dependent
Methionine gamma-lyase (EC 4.4.1.11)
Methionyl-tRNA formyltransferase (EC 2.1.2.9)
Methyl-accepting chemotaxis protein I (serine chemoreceptor protein)
Methylamine dehydrogenase light chain precursor (EC 1.4.99.3)
Methylated-DNA--protein-cysteine methyltransferase (EC 2.1.1.63)
Methylcrotonyl-CoA carboxylase biotin-containing subunit (EC 6.4.1.4)
Methylcrotonyl-CoA carboxylase carboxyl transferase subunit (EC 6.4.1.4)
Methylenetetrahydrofolate dehydrogenase (NADP+) (EC 1.5.1.5)
Methylglutaconyl-CoA hydratase (EC 4.2.1.18)
Methylglyoxal synthase (EC 4.2.3.3)
Methylisocitrate lyase (EC 4.1.3.30)
Methylmalonate-semialdehyde dehydrogenase (EC 1.2.1.27)
Methylmalonate-semialdehyde dehydrogenase [inositol] (EC 1.2.1.27)
Methylmalonyl-CoA mutase (EC 5.4.99.2)
Methylmalonyl-CoA:Pyruvate transcarboxylase 12S subunit (EC 2.1.3.1)
Methylmalonyl-CoA:Pyruvate transcarboxylase 5S subunit (EC 2.1.3.1)
Methylthioadenosine deaminase
Methylthioribose-1-phosphate isomerase (EC 5.3.1.23)
Microsomal dipeptidase (EC 3.4.13.19)
Minor pilin of type IV secretion complex (VirB5)

L-carnitine/gamma-butyrobetaine antiporter
L-cystine uptake protein TcyP
L-fucose isomerase (EC 5.3.1.25)
L-fucose mutarotase
L-fucose operon activator
L-fuculokinase (EC 2.7.1.51)
L-idonate 5-dehydrogenase (EC 1.1.1.264)
L-idonate, D-gluconate, 5-keto-D-gluconate transporter
L-lactate dehydrogenase (EC 1.1.1.27)
L-lactate dehydrogenase (EC 1.1.2.3)
L-lactate permease
L-lysine permease
L-proline glycine betaine ABC transport system permease protein ProV (TC 3.A.1.12.1)
L-proline glycine betaine ABC transport system permease protein ProW (TC 3.A.1.12.1)
L-proline glycine betaine binding ABC transporter protein ProX (TC 3.A.1.12.1)
L-Proline/Glycine betaine transporter ProP
L-rhamnose isomerase (EC 5.3.1.14)
L-rhamnose operon regulatory protein RhaS
L-ribulose-5-phosphate 4-epimerase (EC 5.1.3.4)
L-serine dehydratase (EC 4.3.1.17)
L-serine dehydratase, alpha subunit (EC 4.3.1.17)
L-seryl-tRNA(Sec) selenium transferase (EC 2.9.1.1)
L-threonine 3-dehydrogenase (EC 1.1.1.103)
L-threonine 3-O-phosphate decarboxylase (EC 4.1.1.81)
L-xylulose/3-keto-L-gulonate kinase (EC 2.7.1.-)
L,D-transpeptidase YbiS
L,D-transpeptidase YcbB
L,D-transpeptidase YnhG
Lactaldehyde reductase (EC 1.1.1.77)
Lactate-responsive regulator LldR in Enterobacteria, GntR family

Mn-dependent transcriptional regulator MntR	Lacto-N-biose phosphorylase (EC 2.4.1.211)
Molybdenum ABC transporter, periplasmic molybdenum-binding protein ModA (TC 3.A.1.8.1)	Lactoylglutathione lyase (EC 4.4.1.5)
Molybdenum cofactor biosynthesis protein MoaB	Large exoproteins involved in heme utilization or adhesion
Molybdenum cofactor biosynthesis protein MoaC	Late competence protein ComEC, DNA transport
Molybdenum cofactor biosynthesis protein MoaE	Leader peptidase (Prepilin peptidase) (EC 3.4.23.43)
Molybdenum transport ATP-binding protein ModC (TC 3.A.1.8.1)	Legionaminic acid synthase (EC 2.5.1.56)
Molybdenum transport system permease protein ModB (TC 3.A.1.8.1)	Leucine-responsive regulatory protein, regulator for leucine (or lrp) regulon and high-affinity branched-chain amino acid transport system
Molybdopterin biosynthesis protein MoeA	Leucyl/phenylalanyl-tRNA--protein transferase (EC 2.3.2.6)
Molybdopterin biosynthesis protein MoeB	Light-independent protochlorophyllide reductase subunit N (EC 1.18.-.-)
Molybdopterin-guanine dinucleotide biosynthesis protein MobB	Lipid A biosynthesis lauroyl acyltransferase (EC 2.3.1.-)
Monoamine oxidase (1.4.3.4)	Lipid A export ATP-binding/permease protein MsbA
Monofunctional biosynthetic peptidoglycan transglycosylase (EC 2.4.2.-)	Lipid-A-disaccharide synthase (EC 2.4.1.182)
Monoglyceride lipase (EC 3.1.1.23)	Lipoate synthase
MotA/TolQ/ExbB proton channel family protein	Lipoate-protein ligase A
MSHA biogenesis protein MshO	Lipocalin-related protein and Bos/Can/Equ allergen
Muconate cycloisomerase (EC 5.5.1.1)	Lipopolysaccharide 1,6-galactosyltransferase (EC 2.4.1.-)
Multicopper oxidase	Lipopolysaccharide ABC transporter, ATP-binding protein LptB
Multidrug efflux membrane fusion protein MexE	Lipopolysaccharide biosynthesis protein WzxC
Multidrug efflux transporter MexF	Lipopolysaccharide core biosynthesis glycosyltransferase, group 2 family protein (EC 2.4.1.-)
Multidrug resistance transporter, Bcr/CflA family	Lipopolysaccharide core biosynthesis protein RfaY
Multidrug transporter MdtB	Lipopolysaccharide core biosynthesis protein WaaP (EC 2.7.-.-), heptosyl-I-kinase
Multimodular transpeptidase-transglycosylase (EC 2.4.1.129) (EC 3.4.-.-)	Lipopolysaccharide heptosyltransferase I (EC 2.4.1.-)
Myo-inositol 2-dehydrogenase (EC 1.1.1.18)	Lipopolysaccharide heptosyltransferase III (EC 2.4.1.-)
N-Acetyl-D-glucosamine ABC transport system ATP-binding protein	Lipoprotein signal peptidase (EC 3.4.23.36)
N-Acetyl-D-glucosamine ABC transport system, sugar-binding protein	Long-chain-fatty-acid--CoA ligase (EC 6.2.1.3)
N-acetyl-gamma-glutamyl-phosphate reductase (EC 1.2.1.38)	Low-affinity gluconate/H+ symporter GntU
N-acetyl-L,L-diaminopimelate deacetylase (EC 3.5.1.47)	Low-affinity inorganic phosphate transporter
N-acetylgalactosamine-6-sulfatase (EC 3.1.6.4)	low-specificity D-threonine aldolase

N-acetylglucosamine kinase of eukaryotic type (EC 2.7.1.59)

N-acetylglucosamine related transporter, NagX

N-acetylglucosamine-1-phosphate uridyltransferase (EC 2.7.7.23)

N-acetylglucosamine-6-phosphate deacetylase (EC 3.5.1.25)

N-acetylglucosamine-regulated TonB-dependent outer membrane receptor
N-acetylglutamate synthase (EC 2.3.1.1)
N-acetylmannosamine kinase (EC 2.7.1.60)

N-acetylmannosamine-6-phosphate 2-epimerase (EC 5.1.3.9)

N-Acetylneuraminate cytidylyltransferase (EC 2.7.7.43)
N-acetylneuraminate lyase (EC 4.1.3.3)
N-acylglucosamine 2-epimerase (EC 5.1.3.8)
N-formylglutamate deformylase (EC 3.5.1.68)

N-methylhydantoinase (ATP-hydrolyzing) (EC 3.5.2.14)
N-methylhydantoinase A (EC 3.5.2.14)
N-methylhydantoinase B (EC 3.5.2.14)

N-succinyl-L,L-diaminopimelate aminotransferase alternative (EC 2.6.1.17)

N-succinyl-L,L-diaminopimelate desuccinylase (EC 3.5.1.18)
Na(+) H(+) antiporter subunit A
Na(+) H(+) antiporter subunit A (TC 2.A.63.1.3)
Na(+) H(+) antiporter subunit B
Na(+) H(+) antiporter subunit E
Na(+) H(+) antiporter subunit G (TC 2.A.63.1.3)

Na(+)-translocating NADH-quinone reductase subunit A (EC 1.6.5.-)

Na(+)-translocating NADH-quinone reductase subunit B (EC 1.6.5.-)

Na(+)-translocating NADH-quinone reductase subunit F (EC 1.6.5.-)
Na+/H+ antiporter NhaA type
NAD kinase (EC 2.7.1.23)
NAD synthetase (EC 6.3.1.5)

NAD-dependent formate dehydrogenase alpha subunit

NAD-dependent formate dehydrogenase beta subunit

NAD-dependent glyceraldehyde-3-phosphate dehydrogenase (EC 1.2.1.12)
NAD-dependent malic enzyme (EC 1.1.1.38)

Low-specificity L-threonine aldolase (EC 4.1.2.48)

LPS-assembly lipoprotein RlpB precursor (Rare lipoprotein B)

LptA, protein essential for LPS transport across the periplasm

LrgA-associated membrane protein LrgB

LsrR, transcriptional repressor of lsr operon
LSU m3Psi1915 methyltransferase RlmH
LSU m5C1962 methyltransferase RlmI

LSU ribosomal protein L10p (P0)

LSU ribosomal protein L11p (L12e)
LSU ribosomal protein L13p (L13Ae)
LSU ribosomal protein L14p (L23e)
LSU ribosomal protein L15p (L27Ae)

LSU ribosomal protein L16p (L10e)
LSU ribosomal protein L17p
LSU ribosomal protein L18p (L5e)

LSU ribosomal protein L19p

LSU ribosomal protein L1p (L10Ae)
LSU ribosomal protein L20p
LSU ribosomal protein L21p
LSU ribosomal protein L23p (L23Ae)
LSU ribosomal protein L24p (L26e)
LSU ribosomal protein L25p

LSU ribosomal protein L27p

LSU ribosomal protein L28p

LSU ribosomal protein L2p (L8e)
LSU ribosomal protein L30p (L7e)
LSU ribosomal protein L32p
LSU ribosomal protein L33p, zinc-independent

LSU ribosomal protein L35p

LSU ribosomal protein L3p (L3e)

LSU ribosomal protein L4p (L1e)
LSU ribosomal protein L5p (L11e)

NAD-dependent protein deacetylase of SIR2 family
NAD-specific glutamate dehydrogenase (EC 1.4.1.2)
NAD-specific glutamate dehydrogenase (EC 1.4.1.2), large form
NAD(P) transhydrogenase alpha subunit (EC 1.6.1.2)
NAD(P) transhydrogenase subunit beta (EC 1.6.1.2)
NAD(P)HX dehydratase
NADH dehydrogenase (EC 1.6.99.3)
NADH oxidoreductase hcr (EC 1.-.-.-)
NADH pyrophosphatase (EC 3.6.1.22)
NADH ubiquinone oxidoreductase chain A (EC 1.6.5.3)
NADH-ubiquinone oxidoreductase chain B (EC 1.6.5.3)
NADH-ubiquinone oxidoreductase chain C (EC 1.6.5.3)
NADH-ubiquinone oxidoreductase chain D (EC 1.6.5.3)
NADH-ubiquinone oxidoreductase chain E (EC 1.6.5.3)
NADH-ubiquinone oxidoreductase chain F (EC 1.6.5.3)
NADH-ubiquinone oxidoreductase chain G (EC 1.6.5.3)
NADH-ubiquinone oxidoreductase chain H (EC 1.6.5.3)
NADH-ubiquinone oxidoreductase chain I (EC 1.6.5.3)
NADH-ubiquinone oxidoreductase chain J (EC 1.6.5.3)
NADH-ubiquinone oxidoreductase chain K (EC 1.6.5.3)
NADH-ubiquinone oxidoreductase chain L (EC 1.6.5.3)
NADH-ubiquinone oxidoreductase chain L (EC 1.6.5.3)
NADH-ubiquinone oxidoreductase chain M (EC 1.6.5.3)
NADH-ubiquinone oxidoreductase chain N (EC 1.6.5.3)
NADP-dependent malic enzyme (EC 1.1.1.40)
NADP-specific glutamate dehydrogenase (EC 1.4.1.4)
NADPH dependent preQ0 reductase (EC 1.7.1.13)
NADPH-dependent glyceraldehyde-3-phosphate dehydrogenase (EC 1.2.1.13)
Naphthoate synthase (EC 4.1.3.36)
Negative regulator of allantoin and glyoxylate utilization operons

LSU ribosomal protein L6p (L9e)
LSU ribosomal protein L7/L12 (P1/P2)
LSU ribosomal protein L9p
Lysine decarboxylase (EC 4.1.1.18)
Lysine decarboxylase 2, constitutive (EC 4.1.1.18)
Lysine decarboxylase, inducible (EC 4.1.1.18)
Lysine/cadaverine antiporter membrane protein CadB
Lysophospholipase (EC 3.1.1.5)
Lysophospholipase L2 (EC 3.1.1.5)
Lysophospholipid acyltransferase
LysR family transcriptional regulator lrhA
LysR family transcriptional regulator near succinyl-CoA:3-ketoacid-coenzyme A transferase
LysR family transcriptional regulator STM3121
LysR family transcriptional regulator YbhD
LysR family transcriptional regulator YfeR
LysR family transcriptional regulator YhjC
LysR family transcriptional regulator YhjC
LysR family transcriptional regulator YnfL
Lysyl-tRNA synthetase (class I) (EC 6.1.1.6)
Lysyl-tRNA synthetase (class II) (EC 6.1.1.6)
Lysyl-tRNA synthetase (class II) related protein found fused to membrane protein
Macrolide export ATP-binding/permease protein MacB (EC 3.6.3.-)
Macrolide-specific efflux protein MacA
Magnesium and cobalt efflux protein CorC
Magnesium and cobalt transport protein CorA
Major myo-inositol transporter IolT
Major pilus subunit of type IV secretion complex (VirB2)
Malate dehydrogenase (EC 1.1.1.37)
Malate synthase (EC 2.3.3.9)
Malate synthase G (EC 2.3.3.9)

Neopullulanase (EC 3.2.1.135)
Neuraminidase NanP
NG,NG-dimethylarginine dimethylaminohydrolase 1 (EC 3.5.3.18)
Ni,Fe-hydrogenase I cytochrome b subunit
Niacin transporter NiaP
Nicel/Cobalt-specific TonB-dependent outer membrane receptor
Nickel ABC transporter, periplasmic nickel-binding protein NikA (TC 3.A.1.5.3)
Nickel transport system permease protein NikC (TC 3.A.1.5.3)
Nicotinamidase (EC 3.5.1.19)
Nicotinamidase family protein YcaC
Nicotinamide phosphoribosyltransferase (EC 2.4.2.12)
Nicotinamide-nucleotide adenylyltransferase, NadM family (EC 2.7.7.1)
Nicotinamide-nucleotide adenylyltransferase, NadR family (EC 2.7.7.1)
Nicotinate phosphoribosyltransferase (EC 2.4.2.11)
Nicotinate-nucleotide adenylyltransferase (EC 2.7.7.18)
Nicotinate-nucleotide--dimethylbenzimidazole phosphoribosyltransferase (EC 2.4.2.21)
NifX-associated protein
NifZ protein
Nitrate ABC transporter, nitrate-binding protein
Nitrate ABC transporter, permease protein
Nitrate/nitrite response regulator protein
Nitrate/nitrite transporter
Nitrate/nitrite transporter NarK
Nitric oxide -responding transcriptional regulator NnrR (Crp/Fnr family)
Nitric-oxide reductase (EC 1.7.99.7), quinol-dependent
Nitrilotriacetate monooxygenase component B (EC 1.14.13.-)
Nitrite reductase [NAD(P)H] large subunit (EC 1.7.1.4)
Nitrite reductase [NAD(P)H] small subunit (EC 1.7.1.4)
Nitrite reductase associated c-type cytochorome NirN
Malate:quinone oxidoreductase (EC 1.1.5.4)
Maleate cis-trans isomerase (EC 5.2.1.1)
Maleylacetoacetate isomerase (EC 5.2.1.2)
Malonate decarboxylase alpha subunit
Malonate decarboxylase beta subunit
Malonate decarboxylase delta subunit
Malonate decarboxylase gamma subunit
Malonate transporter, MadM subunit
Malonyl CoA-acyl carrier protein transacylase (EC 2.3.1.39)
Malonyl-CoA decarboxylase (EC 4.1.1.9)
Malto-oligosyltrehalose synthase (EC 5.4.99.15)
Maltodextrin glucosidase (EC 3.2.1.20)
Maltodextrin phosphorylase (EC 2.4.1.1)
Maltoporin (maltose/maltodextrin high-affinity receptor, phage lambda receptor protein)
Maltose O-acetyltransferase (EC 2.3.1.79)
Maltose operon periplasmic protein MalM
Maltose operon transcriptional repressor MalR, LacI family
Maltose phosphorylase (EC 2.4.1.8)
Maltose-6'-phosphate glucosidase (EC 3.2.1.122)
Maltose/maltodextrin ABC transporter, permease protein MalF
Maltose/maltodextrin ABC transporter, permease protein MalG
Maltose/maltodextrin ABC transporter, substrate binding periplasmic protein MalE
Maltose/maltodextrin transport ATP-binding protein MalK (EC 3.6.3.19)
Malyl-CoA lyase (EC 4.1.3.24)
mandelate racemase/muconate lactonizing enzyme family protein
Manganese superoxide dismutase (EC 1.15.1.1)
Manganese transport protein MntH
Manganese-dependent inorganic pyrophosphatase (EC 3.6.1.1)
Mannitol operon repressor

Nitrite transporter from formate/nitrite family

Nitrogen regulation protein NR(I)

Nitrogen regulatory protein P-II

Nitrogenase (molybdenum-iron) beta chain (EC 1.18.6.1)

Nitrogenase (molybdenum-iron) reductase and maturation protein NifH

Nitrogenase FeMo-cofactor scaffold and assembly protein NifE

Nitrogenase FeMo-cofactor synthesis FeS core scaffold and assembly protein NifB

Nitrogenase FeMo-cofactor synthesis molybdenum delivery protein NifQ

Nitrous oxide reductase maturation protein NosD

Nitrous oxide reductase maturation protein NosF (ATPase)
NnrS protein involved in response to NO

Non-ribosomal peptide synthetase modules, pyoverdine??
Non-specific DNA-binding protein Dps

NrfD protein

Nucleoside 5-triphosphatase RdgB (dHAPTP, dITP, XTP-specific) (EC 3.6.1.15)

Nucleoside diphosphate kinase (EC 2.7.4.6)

Nucleoside permease NupC

Nucleoside triphosphate pyrophosphohydrolase MazG (EC 3.6.1.8)

Nucleoside-specific channel-forming protein Tsx precursor

Nucleotidase YfbR, HD superfamily

Nucleotide excision repair protein, with UvrB/UvrC motif
Nucleotide pyrophosphatase (EC 3.6.1.9)
Nudix hydrolase family protein YffH

Nudix-like NDP and NTP phosphohydrolase YmfB

O-acetylhomoserine sulfhydrylase (EC 2.5.1.49)

O-succinylbenzoic acid--CoA ligase (EC 6.2.1.26)

O-succinylhomoserine sulfhydrylase (EC 2.5.1.48)

Octanoate-[acyl-carrier-protein]-protein-N-octanoyltransferase

Oligopeptidase A (EC 3.4.24.70)
Oligopeptide ABC transporter, periplasmic oligopeptide-binding protein OppA

Mannonate dehydratase (EC 4.2.1.8)

Mannose-1-phosphate guanylyltransferase (EC 2.7.7.13)

Mannose-1-phosphate guanylyltransferase (GDP) (EC 2.7.7.22)

Mannose-6-phosphate isomerase (EC 5.3.1.8)

mannose-specific adhesin FimH

Mediator of hyperadherence YidE

Melibiose carrier protein, Na+/melibiose symporter

membrane c-type cytochrome cy

Membrane fusion component of tripartite multidrug resistance system

Membrane Protein Functionally coupled to the MukBEF Chromosome Partitioning Mechanism
Membrane protein YcjF

Membrane transporter HdeD, H-NS repressed
Membrane-associated zinc metalloprotease

Membrane-bound lytic murein transglycosylase B (EC 3.2.1.-)

Membrane-bound lytic murein transglycosylase B precursor (EC 3.2.1.-)

Membrane-bound lytic murein transglycosylase D precursor (EC 3.2.1.-)

Membrane-bound lytic murein transglycosylase E (EC 3.2.1.-)

Menaquinone via futalosine step 1

Menaquinone via futalosine step 3

Menaquinone-specific isochorismate synthase (EC 5.4.4.2)

Mercuric ion reductase (EC 1.16.1.1)
Mercuric resistance operon regulatory protein
Mercuric transport protein, MerT

Metal-dependent hydrolase involved in phosphonate metabolism

Metal-dependent hydrolase YbeY, involved in rRNA and/or ribosome maturation and assembly

Metal-dependent hydrolases of the beta-lactamase superfamily I

Methane monooxygenase component A beta chain (EC 1.14.13.25)

Methanol dehydrogenase large subunit protein (EC 1.1.99.8)

Methenyltetrahydrofolate cyclohydrolase (EC 3.5.4.9)
Methionine ABC transporter ATP-binding protein

(TC 3.A.1.5.1)

Oligopeptide transport ATP-binding protein OppD (TC 3.A.1.5.1)

Oligopeptide transport ATP-binding protein OppF (TC 3.A.1.5.1)

Oligopeptide transport system permease protein OppB (TC 3.A.1.5.1)

Oligopeptide transport system permease protein OppC (TC 3.A.1.5.1)

Omega-amino acid--pyruvate aminotransferase (EC 2.6.1.18)

Optional hypothetical component of the B12 transporter BtuN

Organic hydroperoxide resistance protein

Organosulfonate utilization protein SsuF

Ornithine aminotransferase (EC 2.6.1.13)

Ornithine carbamoyltransferase (EC 2.1.3.3)

Ornithine cyclodeaminase (EC 4.3.1.12)

Ornithine decarboxylase (EC 4.1.1.17)

Orotate phosphoribosyltransferase (EC 2.4.2.10)
Orotidine 5'-phosphate decarboxylase (EC 4.1.1.23)
Osmolarity sensory histidine kinase EnvZ

Osmosensitive K+ channel histidine kinase KdpD (EC 2.7.3.-)
Osmotically activated L-carnitine/choline ABC transporter, ATP-binding protein OpuCA
Osmotically activated L-carnitine/choline ABC transporter, permease protein OpuCD

Outer membrane component of tripartite multidrug resistance system

Outer membrane lipoprotein omp16 precursor
Outer membrane porin, OprD family

Outer membrane protein A precursor
Outer membrane protein C precursor
Outer membrane protein OprN

outer membrane protein, MtrB

Outer membrane stress sensor protease DegQ, serine protease

Outer membrane stress sensor protease DegS
Outer membrane usher protein HtrE

Methionine ABC transporter permease protein

Methionine ABC transporter substrate-binding protein

Methionine aminotransferase, PLP-dependent

Methionine transporter MetT

Methionyl-tRNA formyltransferase (EC 2.1.2.9)

Methoxyneurosporene dehydrogenase (EC 1.14.99.-)

Methyl-accepting chemotaxis protein I (serine chemoreceptor protein)

Methyl-accepting chemotaxis protein II (aspartate chemoreceptor protein)

Methyl-accepting chemotaxis protein III (ribose and galactose chemoreceptor protein)

Methylated-DNA--protein-cysteine methyltransferase (EC 2.1.1.63)

Methylcrotonyl-CoA carboxylase biotin-containing subunit (EC 6.4.1.4)

Methylcrotonyl-CoA carboxylase carboxyl transferase subunit (EC 6.4.1.4)

Methylenetetrahydrofolate dehydrogenase (NADP+) (EC 1.5.1.5)
Methylglutaconyl-CoA hydratase (EC 4.2.1.18)
Methylglyoxal synthase (EC 4.2.3.3)

Methylisocitrate lyase (EC 4.1.3.30)

Methylmalonate-semialdehyde dehydrogenase (EC 1.2.1.27)

Methylmalonate-semialdehyde dehydrogenase [inositol] (EC 1.2.1.27)

Methylmalonyl-CoA mutase (EC 5.4.99.2)

Methylmalonyl-CoA:Pyruvate transcarboxylase 5S subunit (EC 2.1.3.1)
Methylthioadenosine deaminase

Methylthioribose-1-phosphate isomerase (EC 5.3.1.23)
Mhp operon transcriptional activator
Mhp operon transcriptional activator
Mlc, transcriptional repressor of MalT (the transcriptional activator of maltose regulon) and manXYZ operon

Molybdate-binding domain of ModE

Molybdenum ABC transporter, periplasmic molybdenum-binding protein ModA (TC 3.A.1.8.1)
Molybdenum cofactor biosynthesis protein MoaB

Outer membrane vitamin B12 receptor BtuB
Oxalyl-CoA decarboxylase (EC 4.1.1.8)
oxidoreductase of aldo/keto reductase family, subgroup 1
P-hydroxybenzoate hydroxylase (EC 1.14.13.2)
P-hydroxylaminobenzoate lyase
Pantoate--beta-alanine ligase (EC 6.3.2.1)
Pantothenate kinase (EC 2.7.1.33)
Pantothenate kinase type II, eukaryotic (EC 2.7.1.33)
Pantothenate kinase type III, CoaX-like (EC 2.7.1.33)
Para-aminobenzoate synthase, amidotransferase component (EC 2.6.1.85)
Para-aminobenzoate synthase, aminase component (EC 2.6.1.85)
Paraquat-inducible protein A
Paraquat-inducible protein B
PdxA-like
Penicillin-binding protein 2 (PBP-2)
Penicillin-insensitive transglycosylase (EC 2.4.2.-) & transpeptidase PBP-1C
Peptide chain release factor 1
Peptide chain release factor 2
Peptide deformylase (EC 3.5.1.88)
Peptide methionine sulfoxide reductase MsrA (EC 1.8.4.11)
Peptide transport system ATP-binding protein SapF
Peptidoglycan-associated lipoprotein precursor
Peptidyl-prolyl cis-trans isomerase (EC 5.2.1.8)
Peptidyl-prolyl cis-trans isomerase ppiA precursor (EC 5.2.1.8)
Peptidyl-prolyl cis-trans isomerase ppiC (EC 5.2.1.8)
Peptidyl-prolyl cis-trans isomerase ppiD (EC 5.2.1.8)
Peptidyl-tRNA hydrolase (EC 3.1.1.29)
Periplasmic alpha-amylase (EC 3.2.1.1)
Periplasmic aromatic aldehyde oxidoreductase, FAD binding subunit YagS
Periplasmic beta-glucosidase (EC 3.2.1.21)

Molybdenum cofactor biosynthesis protein MoaC
Molybdenum cofactor biosynthesis protein MoaE
Molybdenum transport ATP-binding protein ModC (TC 3.A.1.8.1)
Molybdenum transport system permease protein ModB (TC 3.A.1.8.1)
Molybdenum transport system protein ModD
Molybdopterin biosynthesis Mog protein, molybdochelatase
Molybdopterin biosynthesis protein MoeA
Molybdopterin biosynthesis protein MoeB
Molybdopterin-containing protein YdhV, similar to aldehyde:ferredoxin oxidoreductase
Molybdopterin-guanine dinucleotide biosynthesis protein MobA
Monoamine oxidase (1.4.3.4)
Monofunctional biosynthetic peptidoglycan transglycosylase (EC 2.4.2.-)
Monoglyceride lipase (EC 3.1.1.23)
MotA/TolQ/ExbB proton channel family protein
Muconate cycloisomerase (EC 5.5.1.1)
Multi antimicrobial extrusion protein (Na(+)/drug antiporter), MATE family of MDR efflux pumps
Multicopper oxidase
Multidrug resistance transporter, Bcr/CflA family
Multidrug transporter MdtB
Multidrug-efflux transporter, major facilitator superfamily (MFS) (TC 2.A.1)
Multimodular transpeptidase-transglycosylase (EC 2.4.1.129) (EC 3.4.-.-)
Multiple sugar ABC transporter, ATP-binding protein
Murein-DD-endopeptidase (EC 3.4.99.-)
Mutator mutT protein (7,8-dihydro-8-oxoguanine-triphosphatase) (EC 3.6.1.-)
MutS domain protein, family 5
Myo-inositol 2-dehydrogenase (EC 1.1.1.18)
Myo-inositol 2-dehydrogenase 1 (EC 1.1.1.18)
N-Acetyl-D-glucosamine ABC transport system, permease protein 1
N-Acetyl-D-glucosamine ABC transport system, permease protein 2
N-Acetyl-D-glucosamine ABC transport system, sugar-binding protein

Periplasmic chorismate mutase I precursor (EC 5.4.99.5)

Periplasmic hemin-binding protein
Periplasmic molybdate-binding domain

Periplasmic nitrate reductase precursor (EC 1.7.99.4)

Periplasmic protein p19 involved in high-affinity Fe2+ transport

Periplasmic thiol:disulfide interchange protein DsbA

Permease of the drug/metabolite transporter (DMT) superfamily

Peroxidase (EC 1.11.1.7)

Phage portal protein
Phage tail fiber protein
Phage tape measure

Phenylacetate-coenzyme A ligase (EC 6.2.1.30)

Phenylalanine hydroxylase transcriptional activator PhhR
Phenylalanine-4-hydroxylase (EC 1.14.16.1)
Phenylalanine-specific permease

Phenylalanyl-tRNA synthetase alpha chain (EC 6.1.1.20)

Phenylalanyl-tRNA synthetase beta chain (EC 6.1.1.20)

PhnB protein
PhnG protein
PhnI protein

PhnJ protein

Phosphatase, Ppx/GppA family
Phosphate ABC transporter, periplasmic phosphate-binding protein PstS (TC 3.A.1.7.1)
Phosphate acetyltransferase (EC 2.3.1.8)

Phosphate regulon sensor protein PhoR (SphS) (EC 2.7.13.3)

Phosphate regulon transcriptional regulatory protein PhoB (SphR)

Phosphate starvation-inducible protein PhoH, predicted ATPase

Phosphate transport ATP-binding protein PstB (TC 3.A.1.7.1)

Phosphate transport regulator (distant homolog of PhoU)

N-acetyl-gamma-glutamyl-phosphate reductase (EC 1.2.1.38)

N-acetyl-L,L-diaminopimelate deacetylase (EC 3.5.1.47)
N-acetylgalactosamine-6-sulfatase (EC 3.1.6.4)

N-acetylglucosamine kinase of eukaryotic type (EC 2.7.1.59)

N-acetylglucosamine-1-phosphate uridyltransferase (EC 2.7.7.23)

N-acetylglucosamine-6-phosphate deacetylase (EC 3.5.1.25)

N-acetylglucosamine-6P-responsive transcriptional repressor NagC, ROK family

N-acetylglucosamine-regulated outer membrane porin

N-acetylglucosamine-regulated TonB-dependent outer membrane receptor
N-acetylglutamate synthase (EC 2.3.1.1)
N-acetylmannosamine kinase (EC 2.7.1.60)

N-acetylmannosamine-6-phosphate 2-epimerase (EC 5.1.3.9)

N-acetylmannosaminyltransferase (EC 2.4.1.187)
N-acetylmuramic acid 6-phosphate etherase
N-acyl homoserine lactone hydrolase

N-acylglucosamine 2-epimerase (EC 5.1.3.8)

N-formylglutamate deformylase (EC 3.5.1.68)

N-formylglutamate deformylase (EC 3.5.1.68) [alternative form]
N-methylhydantoinase A (EC 3.5.2.14)
N-methylhydantoinase B (EC 3.5.2.14)

N-succinyl-L,L-diaminopimelate aminotransferase alternative (EC 2.6.1.17)

N-succinyl-L,L-diaminopimelate desuccinylase (EC 3.5.1.18)

Na+/H+ antiporter NhaA type
NAD kinase (EC 2.7.1.23)

NAD synthetase (EC 6.3.1.5)

NAD-dependent formate dehydrogenase alpha subunit

NAD-dependent formate dehydrogenase beta subunit

NAD-dependent formate dehydrogenase gamma subunit

NAD-dependent glyceraldehyde-3-phosphate dehydrogenase (EC 1.2.1.12)

Phosphate transport system permease protein PstA (TC 3.A.1.7.1)

Phosphate transport system permease protein PstC (TC 3.A.1.7.1)

Phosphate transport system regulatory protein PhoU

Phosphate:acyl-ACP acyltransferase PlsX

Phosphatidate cytidylyltransferase (EC 2.7.7.41)
Phosphatidylglycerophosphatase A (EC 3.1.3.27)

Phosphatidylglycerophosphatase B (EC 3.1.3.27)
Phosphatidylserine decarboxylase (EC 4.1.1.65)

Phospho-N-acetylmuramoyl-pentapeptide-transferase (EC 2.7.8.13)

Phosphoadenylyl-sulfate reductase [thioredoxin] (EC 1.8.4.8)

Phosphocarrier protein, nitrogen regulation associated

Phosphoenolpyruvate carboxykinase [ATP] (EC 4.1.1.49)

Phosphoenolpyruvate carboxykinase [GTP] (EC 4.1.1.32)

Phosphoenolpyruvate carboxylase (EC 4.1.1.31)

Phosphoenolpyruvate synthase (EC 2.7.9.2)
Phosphoenolpyruvate-dihydroxyacetone phosphotransferase (EC 2.7.1.121), dihydroxyacetone binding subunit DhaK
Phosphoenolpyruvate-dihydroxyacetone phosphotransferase (EC 2.7.1.121), subunit DhaM
Phosphoenolpyruvate-dihydroxyacetone phosphotransferase operon regulatory protein DhaR

Phosphoenolpyruvate-protein phosphotransferase of PTS system (EC 2.7.3.9)

Phosphoesterase, DHH family protein
Phosphoethanolamine transferase specific for the outer Kdo residue of lipopolysaccharide

Phosphogluconate dehydratase (EC 4.2.1.12)

Phosphoglucosamine mutase (EC 5.4.2.10)

Phosphoglycerate kinase (EC 2.7.2.3)

Phosphoglycolate phosphatase (EC 3.1.3.18)
Phosphoheptose isomerase (EC 5.3.1.-)
Phosphoheptose isomerase 1 (EC 5.3.1.-)

NAD-dependent malic enzyme (EC 1.1.1.38)

NAD-dependent protein deacetylase of SIR2 family

NAD-specific glutamate dehydrogenase (EC 1.4.1.2)

NAD-specific glutamate dehydrogenase (EC 1.4.1.2), large form

NAD(P) transhydrogenase alpha subunit (EC 1.6.1.2)
NAD(P) transhydrogenase subunit beta (EC 1.6.1.2)

NAD(P)H-flavin reductase (EC 1.5.1.29) (EC 1.16.1.3)
NAD(P)HX epimerase

NADH dehydrogenase (EC 1.6.99.3)

NADH dehydrogenase, subunit 5

NADH pyrophosphatase (EC 3.6.1.22)

NADH ubiquinone oxidoreductase chain A (EC 1.6.5.3)

NADH-ubiquinone oxidoreductase chain B (EC 1.6.5.3)

NADH-ubiquinone oxidoreductase chain C (EC 1.6.5.3)

NADH-ubiquinone oxidoreductase chain D (EC 1.6.5.3)

NADH-ubiquinone oxidoreductase chain E (EC 1.6.5.3)

NADH-ubiquinone oxidoreductase chain F (EC 1.6.5.3)

NADH-ubiquinone oxidoreductase chain G (EC 1.6.5.3)

NADH-ubiquinone oxidoreductase chain H (EC 1.6.5.3)

NADH-ubiquinone oxidoreductase chain I (EC 1.6.5.3)

NADH-ubiquinone oxidoreductase chain J (EC 1.6.5.3)

NADH-ubiquinone oxidoreductase chain K (EC 1.6.5.3)

NADH-ubiquinone oxidoreductase chain L (EC 1.6.5.3)

NADH-ubiquinone oxidoreductase chain M (EC 1.6.5.3)

NADH-ubiquinone oxidoreductase chain N (EC 1.6.5.3)
NADH:ubiquinone oxidoreductase 17.2 kD subunit
NADP-dependent malic enzyme (EC 1.1.1.40)

Phospholipid-lipopolysaccharide ABC transporter
Phosphomannomutase (EC 5.4.2.8)
Phosphonate ABC transporter ATP-binding protein (TC 3.A.1.9.1)
Phosphonate ABC transporter permease protein phnE (TC 3.A.1.9.1)
Phosphonate ABC transporter permease protein phnE1 (TC 3.A.1.9.1)
Phosphonate ABC transporter permease protein phnE2 (TC 3.A.1.9.1)
Phosphonate ABC transporter phosphate-binding periplasmic component (TC 3.A.1.9.1)
Phosphonopyruvate decarboxylase (EC 4.1.1.82)
Phosphopantetheine adenylyltransferase (EC 2.7.7.3)
Phosphopantothenoylcysteine synthetase (EC 6.3.2.5)
Phosphoribosyl-AMP cyclohydrolase (EC 3.5.4.19)
Phosphoribosyl-ATP pyrophosphatase (EC 3.6.1.31)
Phosphoribosylamine--glycine ligase (EC 6.3.4.13)
Phosphoribosylaminoimidazole carboxylase ATPase subunit (EC 4.1.1.21)
Phosphoribosylaminoimidazole carboxylase catalytic subunit (EC 4.1.1.21)
Phosphoribosylaminoimidazole-succinocarboxamide synthase (EC 6.3.2.6)
Phosphoribosylaminoimidazolecarboxamide formyltransferase (EC 2.1.2.3)
Phosphoribosylformimino-5-aminoimidazole carboxamide ribotide isomerase (EC 5.3.1.16)
Phosphoribosylformylglycinamidine cyclo-ligase (EC 6.3.3.1)
Phosphoribosylformylglycinamidine synthase, glutamine amidotransferase subunit (EC 6.3.5.3)
Phosphoribosylformylglycinamidine synthase, synthetase subunit (EC 6.3.5.3)
Phosphoribosylglycinamide formyltransferase (EC 2.1.2.2)
Phosphoribosylglycinamide formyltransferase 2 (EC 2.1.2.-)
Phosphoribulokinase (EC 2.7.1.19)
Phosphoribulokinase (EC 2.7.1.19) homolog, function unknown
Phosphoserine aminotransferase (EC 2.6.1.52)
Phosphoserine phosphatase (EC 3.1.3.3)
Phytochrome, two-component sensor histidine kinase (EC 2.7.3.-)

NADP-specific glutamate dehydrogenase (EC 1.4.1.4)
NADPH dependent preQ0 reductase (EC 1.7.1.13)
Naphthoate synthase (EC 4.1.3.36)
Negative regulator of allantoin and glyoxylate utilization operons
Neuraminidase NanP
NG,NG-dimethylarginine dimethylaminohydrolase 1 (EC 3.5.3.18)
Ni,Fe-hydrogenase III small subunit
Ni/Fe-hydrogenase 2 B-type cytochrome subunit
Niacin transporter NiaP
Nicel/Cobalt-specific TonB-dependent outer membrane receptor
Nickel responsive regulator NikR
Nickel transport ATP-binding protein NikD (TC 3.A.1.5.3)
Nickel transport system permease protein NikB (TC 3.A.1.5.3)
Nicotinamidase (EC 3.5.1.19)
Nicotinamidase/isochorismatase family protein
Nicotinamide phosphoribosyltransferase (EC 2.4.2.12)
Nicotinamide-nucleotide adenylyltransferase, NadR family (EC 2.7.7.1)
Nicotinate phosphoribosyltransferase (EC 2.4.2.11)
Nicotinate-nucleotide adenylyltransferase (EC 2.7.7.18)
Nicotinate-nucleotide--dimethylbenzimidazole phosphoribosyltransferase (EC 2.4.2.21)
NifT protein
Nitrate ABC transporter, ATP-binding protein
Nitrate ABC transporter, nitrate-binding protein
Nitrate ABC transporter, permease protein
Nitrate reductase cytochrome c550-type subunit
Nitrate/nitrite response regulator protein
Nitrate/nitrite sensor protein (EC 2.7.3.-)
Nitrate/nitrite transporter

Phytoene dehydrogenase and related proteins
Phytoene desaturase (EC 1.14.99.-)
Phytoene synthase (EC 2.5.1.32)
Plasmid replication protein RepA
Plasmid replication protein RepB
Plasmid replication protein RepC
Poly(3-hydroxyalkanoate) depolymerase
Poly(A) polymerase (EC 2.7.7.19)
Poly(glycerophosphate chain) D-alanine transfer protein DltD
Polygalacturonase (EC 3.2.1.15)
Polyhydroxyalkanoic acid synthase
Polymyxin resistance protein ArnT, undecaprenyl phosphate-alpha-L-Ara4N transferase
Polyphosphate kinase (EC 2.7.4.1)
Polyribonucleotide nucleotidyltransferase (EC 2.7.7.8)
Porphobilinogen synthase (EC 4.2.1.24)
Positive regulator of CheA protein activity (CheW)
Positive transcription regulator EvgA
Possible GPH family transporter (TC 2.A.2) for arabinosides
Possible hypoxanthine oxidase XdhD (EC 1.-.-.-)
Potassium channel protein
Potassium efflux system KefA protein
Potassium uptake protein TrkH
Potassium uptake protein, integral membrane component, KtrB
Potassium-transporting ATPase A chain (EC 3.6.3.12) (TC 3.A.3.7.1)
Potassium-transporting ATPase B chain (EC 3.6.3.12) (TC 3.A.3.7.1)
Potassium-transporting ATPase C chain (EC 3.6.3.12) (TC 3.A.3.7.1)
POTASSIUM/PROTON ANTIPORTER ROSB
Precorrin-6A synthase (EC 2.1.1.152)
Predicted 2-keto-4-pentenoate hydratase/2-oxohepta-3-ene-1,7-dioic acid hydratase

Nitric oxide reductase activation protein NorQ
Nitric-oxide reductase (EC 1.7.99.7), quinol-dependent
Nitrilotriacetate monooxygenase component B (EC 1.14.13.-)
Nitrite reductase [NAD(P)H] large subunit (EC 1.7.1.4)
Nitrite reductase [NAD(P)H] small subunit (EC 1.7.1.4)
Nitrite reductase probable [NAD(P)H] subunit (EC 1.7.1.4)
Nitrite transporter from formate/nitrite family
Nitrogen regulation protein NR(I)
Nitrogen regulatory protein P-II
Nitrogenase (molybdenum-iron) alpha chain (EC 1.18.6.1)
Nitrogenase (molybdenum-iron) beta chain (EC 1.18.6.1)
Nitrogenase (molybdenum-iron) reductase and maturation protein NifH
Nitrogenase (molybdenum-iron)-specific transcriptional regulator NifA
Nitrogenase FeMo-cofactor carrier protein NifX
Nitrogenase FeMo-cofactor scaffold and assembly protein NifE
Nitrogenase FeMo-cofactor scaffold and assembly protein NifN
Nitrogenase FeMo-cofactor synthesis FeS core scaffold and assembly protein NifB
Nitrogenase stabilizing/protective protein NifW
Nitrous oxide reductase maturation protein NosD
Nitrous oxide reductase maturation protein NosF (ATPase)
Nitrous oxide reductase maturation protein NosR
Nitrous-oxide reductase (EC 1.7.99.6)
NMN phosphatase (EC 3.1.3.5)
NnrS protein involved in response to NO
Non-specific DNA-binding protein Dps
Nondeblocking aminopeptidase YpdE (X-X-[^PR]- specific)
Novel Xylose regulator from LacI family
NrfD protein
Nucleoside 5-triphosphatase RdgB (dHAPTP, dITP, XTP-specific) (EC 3.6.1.15)

Predicted ATP-dependent endonuclease of the OLD family, YbjD subgroup

Predicted ATPase related to phosphate starvation-inducible protein PhoH

Predicted beta-glucoside-regulated ABC transport system, permease component 1, COG1175

Predicted beta-glucoside-regulated ABC transport system, sugar binding component, COG1653

Predicted cell-wall-anchored protein SasC (LPXTG motif)

Predicted chaperone lipoprotein YacC, potentially involved in protein secretion

Predicted D-glucarate or D-galactorate regulator, GntR family

Predicted D-lactate dehydrogenase, Fe-S protein, FAD/FMN-containing

Predicted galacto-N-biose-/lacto-N-biose I ABC transporter, permease component 1

Predicted galacto-N-biose-/lacto-N-biose I ABC transporter, permease component 2

Predicted glycogen synthase, ADP-glucose transglucosylase (EC 2.4.1.21), Actinobacterial type

Predicted hydrolase of the metallo-beta-lactamase superfamily, clustered with KDO2-Lipid A biosynthesis genes

Predicted hydroxymethylpyrimidine transporter CytX

Predicted L-lactate dehydrogenase, Fe-S oxidoreductase subunit YkgE

Predicted L-lactate dehydrogenase, hypothetical protein subunit YkgG

Predicted L-lactate dehydrogenase, Iron-sulfur cluster-binding subunit YkgF

Predicted L-rhamnose ABC transporter, substrate-binding component

Predicted lactaldehyde dehydrogenase (EC 1.2.1.22)

Predicted monooxygenase RutA in novel pyrimidine catabolism pathway

Predicted P-loop ATPase fused to an acetyltransferase COG1444

Predicted PTS system, galactosamine-specific IIC component (EC 2.7.1.69)

Predicted regulator PutR for proline utilization, GntR family

Predicted signal-transduction protein containing cAMP-binding and CBS domains

Predicted transcriptional regulator LiuR of leucine degradation pathway, MerR family

Predicted transcriptional regulator of N-Acetylglucosamine utilization, GntR

Nucleoside diphosphate kinase (EC 2.7.4.6)

Nucleoside permease NupC

Nucleoside triphosphate pyrophosphohydrolase MazG (EC 3.6.1.8)

Nucleoside-specific channel-forming protein Tsx precursor

Nucleotidase YfbR, HD superfamily

Nucleotide pyrophosphatase (EC 3.6.1.9)

Nudix hydrolase family protein YffH

O-succinylbenzoate synthase (EC 4.2.1.113)

O-succinylbenzoic acid--CoA ligase (EC 6.2.1.26)

O-succinylhomoserine sulfhydrylase (EC 2.5.1.48)

Octanoate-[acyl-carrier-protein]-protein-N-octanoyltransferase

Oligopeptidase A (EC 3.4.24.70)

Oligopeptide ABC transporter, periplasmic oligopeptide-binding protein OppA (TC 3.A.1.5.1)

Oligopeptide transport ATP-binding protein OppD (TC 3.A.1.5.1)

Oligopeptide transport ATP-binding protein OppF (TC 3.A.1.5.1)

Oligopeptide transport system permease protein OppB (TC 3.A.1.5.1)

Oligopeptide transport system permease protein OppC (TC 3.A.1.5.1)

Omega-amino acid--pyruvate aminotransferase (EC 2.6.1.18)

OpcA, an allosteric effector of glucose-6-phosphate dehydrogenase, actinobacterial

Organic hydroperoxide resistance protein

Organic hydroperoxide resistance transcriptional regulator

Ornithine aminotransferase (EC 2.6.1.13)

Ornithine carbamoyltransferase (EC 2.1.3.3)

Ornithine decarboxylase (EC 4.1.1.17)

Orotate phosphoribosyltransferase (EC 2.4.2.10) family

Predicted transcriptional regulator of pyridoxine metabolism

Predicted transcriptional regulator of the myo-inositol catabolic operon

Predicted trehalose permease, MFS family

Prephenate dehydratase (EC 4.2.1.51)

Prephenate dehydrogenase (EC 1.3.1.12)

Probable 3-phenylpropionic acid transporter

Probable acyl-CoA dehydrogenase (EC 1.3.99.3)

Probable Co/Zn/Cd efflux system membrane fusion protein

Probable glutathione S-transferase (EC 2.5.1.18), YfcG homolog

probable iron binding protein from the HesB_IscA_SufA family

Probable low-affinity inorganic phosphate transporter

probable RuBisCo-expression protein CbbX

Probable transcription regulator protein of MDR efflux pump cluster

Probable transcriptional activator for leuABCD operon

Probable VANILLIN dehydrogenase oxidoreductase protein (EC 1.-.-.-)

Proline iminopeptidase (EC 3.4.11.5)

Propionate catabolism operon regulatory protein PrpR

Propionate catabolism operon transcriptional regulator of GntR family [predicted]

Propionate--CoA ligase (EC 6.2.1.17)

Propionyl-CoA carboxylase beta chain (EC 6.4.1.3)

Propionyl-CoA carboxylase carboxyl transferase subunit (EC 6.4.1.3)

Protease VII (Omptin) precursor (EC 3.4.23.49)

Proteasome subunit beta (EC 3.4.25.1), bacterial

Protein acetyltransferase

Protein AraJ precursor

Protein involved in catabolism of external DNA

Protein of unknown function DUF1009 clustered with KDO2-Lipid A biosynthesis genes

Protein of unknown function DUF1446

Protein of unknown function DUF374

Orotidine 5'-phosphate decarboxylase (EC 4.1.1.23)

Ortho-halobenzoate 1,2-dioxygenase alpha-ISP protein OhbB

Osmolarity sensory histidine kinase EnvZ

Osmoprotectant ABC transporter permease protein YehY

Osmosensitive K+ channel histidine kinase KdpD (EC 2.7.3.-)

Outer membrane component of tripartite multidrug resistance system

Outer membrane lipoprotein omp16 precursor

Outer membrane porin, OprD family

Outer membrane protein A precursor

Outer membrane protein NlpB, lipoprotein component of the protein assembly complex (forms a complex with YaeT, YfiO, and YfgL)

Outer membrane receptor proteins, mostly Fe transport

Outer membrane stress sensor protease DegQ, serine protease

Outer membrane stress sensor protease DegS

Outer membrane usher protein HtrE

Outer membrane vitamin B12 receptor BtuB

Oxalyl-CoA decarboxylase (EC 4.1.1.8)

oxidoreductase of aldo/keto reductase family, subgroup 1

P-hydroxybenzoate hydroxylase (EC 1.14.13.2)

Pantoate--beta-alanine ligase (EC 6.3.2.1)

Pantothenate kinase (EC 2.7.1.33)

Pantothenate kinase type II, eukaryotic (EC 2.7.1.33)

Pantothenate:Na+ symporter (TC 2.A.21.1.1)

Para-aminobenzoate synthase, amidotransferase component (EC 2.6.1.85)

Para-aminobenzoate synthase, aminase component (EC 2.6.1.85)

Paralog of coenzyme PQQ synthesis protein C

Paraquat-inducible protein A

Paraquat-inducible protein B

PdxA-like

Pectinesterase (EC 3.1.1.11)

Protein RcsF	Penicillin-binding protein 2 (PBP-2)
Protein RtcB	Penicillin-binding protein AmpH
Protein YigP (COG3165) clustered with ubiquinone biosynthetic genes	Penicillin-insensitive transglycosylase (EC 2.4.2.-) & transpeptidase PBP-1C
Protein-L-isoaspartate O-methyltransferase (EC 2.1.1.77)	Peptide chain release factor 1
Protein-N(5)-glutamine methyltransferase PrmC, methylates polypeptide chain release factors RF1 and RF2	Peptide chain release factor 2
Protocatechuate 3,4-dioxygenase beta chain (EC 1.13.11.3)	Peptide deformylase (EC 3.5.1.88)
Protocatechuate 4,5-dioxygenase alpha chain (EC 1.13.11.8)	Peptide methionine sulfoxide reductase MsrA (EC 1.8.4.11)
Protocatechuate 4,5-dioxygenase beta chain (EC 1.13.11.8)	Peptide transport periplasmic protein SapA
PTS system, beta-glucoside-specific IIB component (EC 2.7.1.69)	Peptide transport system ATP-binding protein SapF
PTS system, cellobiose-specific IIC component (EC 2.7.1.69)	Peptide transport system permease protein SapC
PTS system, fructose-specific IIB component (EC 2.7.1.69)	Peptidoglycan-associated lipoprotein precursor
PTS system, galactitol-specific IIC component (EC 2.7.1.69)	Peptidyl-prolyl cis-trans isomerase (EC 5.2.1.8)
PTS system, glucitol/sorbitol-specific IIB component and second of two IIC components (EC 2.7.1.69)	Peptidyl-prolyl cis-trans isomerase PpiB (EC 5.2.1.8)
PTS system, mannitol-specific IIB component (EC 2.7.1.69)	Peptidyl-prolyl cis-trans isomerase ppiD (EC 5.2.1.8)
PTS system, N-acetylglucosamine-specific IIB component (EC 2.7.1.69)	Peptidyl-tRNA hydrolase (EC 3.1.1.29)
PTS system, sucrose-specific IIB component (EC 2.7.1.69)	Periplasmic alpha-amylase (EC 3.2.1.1)
Purine nucleoside phosphorylase (EC 2.4.2.1)	Periplasmic aromatic aldehyde oxidoreductase, FAD binding subunit YagS
PurR: transcription regulator associated with purine metabolism	Periplasmic aromatic aldehyde oxidoreductase, iron-sulfur subunit YagT
Putative 5-carboxymethyl-2-hydroxymuconate semialdehyde dehydrogenase oxidoreductase protein (EC 1.2.1.60)	Periplasmic aromatic amino acid aminotransferase beta precursor (EC 2.6.1.57)
Putative conjugative transposon mobilization protein BF0132	Periplasmic beta-glucosidase (EC 3.2.1.21)
putative Cytochrome bd2, subunit I	Periplasmic hemin-binding protein
putative Cytochrome bd2, subunit II	Periplasmic molybdate-binding domain
Putative cytochrome C-type biogenesis protein	Periplasmic nitrate reductase precursor (EC 1.7.99.4)
putative cytochrome P450 hydroxylase	Periplasmic thiol:disulfide interchange protein DsbA
Putative deoxyribonuclease YcfH	Periplasmic thiol:disulfide oxidoreductase DsbB, required for DsbA reoxidation
Putative deoxyribonuclease YjjV	Permease of the drug/metabolite transporter (DMT) superfamily
Putative deoxyribose-specific ABC transporter, ATP-binding protein	Peroxidase (EC 1.11.1.7)
Putative deoxyribose-specific ABC transporter, permease protein	Petrobactin ABC transporter, periplasmic binding protein

Putative diheme cytochrome c-553
Putative Dihydrolipoamide dehydrogenase (EC 1.8.1.4)
Putative dihydroxyacetone kinase (EC 2.7.1.29), dihydroxyacetone binding subunit
Putative formate dehydrogenase oxidoreductase protein
Putative formate dehydrogenase-specific chaperone
Putative fumarylacetoacetate (FAA) hydrolase
Putative glutathione transporter, ATP-binding component
Putative glutathione transporter, permease component
Putative glycosyl/glycerophosphate transferase in teichoic acid biosynthesis
Putative hemin-binding lipoprotein
Putative iron-sulfur cluster assembly scaffold protein for SUF system, SufE2
Putative lipoprotein in cluster with COG2110
Putative lipoprotein, specific for Pseudomonas, in cluster with COG2110
Putative metal chaperone, involved in Zn homeostasis, GTPase of COG0523 family
Putative mobilization protein BF0133
Putative OMR family iron-siderophore receptor precursor
Putative oxidoreductase in putrescine utilization cluster
Putative oxidoreductase YncB
Putative phosphatase YigL, haloacid dehalogenase-like phosphatase family
Putative phosphoenolpyruvate synthase/pyruvate phosphate dikinase, C-terminal domain
Putative preQ0 transporter
Putative sensory histidine kinase YfhA
putative sodium-dependent bicarbonate transporter
Putative sucrose phosphorylase (EC 2.4.1.7)
Putative two-component response regulator and GGDEF family protein YeaJ
Putative two-domain glycosyltransferase
PutR, transcriptional activator of PutA and PutP
Putrescine ABC transporter putrescine-binding protein PotF (TC 3.A.1.11.2)
PF00070 family, FAD-dependent NAD(P)-disulphide oxidoreductase
Phage portal (connector) protein
Phage portal protein
Phage shock protein A
Phage tail fiber protein
Phenol hydroxylase, P4 oxygenase component DmpO (EC 1.14.13.7)
Phenylacetate-CoA oxygenase, PaaI subunit
Phenylacetate-coenzyme A ligase (EC 6.2.1.30) PaaF
Phenylacetic acid degradation protein PaaN, ring-opening aldehyde dehydrogenase (EC 1.2.1.3)
Phenylalanine-4-hydroxylase (EC 1.14.16.1)
Phenylalanine-specific permease
Phenylalanyl-tRNA synthetase alpha chain (EC 6.1.1.20)
Phenylalanyl-tRNA synthetase beta chain (EC 6.1.1.20)
PhnB protein
PhnH protein
PhnI protein
PhnJ protein
Phosphatase, Ppx/GppA family
Phosphate ABC transporter, periplasmic phosphate-binding protein PstS (TC 3.A.1.7.1)
Phosphate acetyltransferase (EC 2.3.1.8)
Phosphate regulon sensor protein PhoR (SphS) (EC 2.7.13.3)
Phosphate regulon transcriptional regulatory protein PhoB (SphR)
Phosphate starvation-inducible protein PhoH, predicted ATPase
Phosphate transport ATP-binding protein PstB (TC 3.A.1.7.1)
Phosphate transport regulator (distant homolog of PhoU)
Phosphate transport system permease protein PstA (TC 3.A.1.7.1)
Phosphate transport system permease protein PstC (TC 3.A.1.7.1)
Phosphate transport system regulatory protein PhoU

Putrescine importer

Putrescine transport ATP-binding protein PotA (TC 3.A.1.11.1)

Putrescine transport ATP-binding protein PotG (TC 3.A.1.11.2)

Putrescine transport system permease protein PotH (TC 3.A.1.11.2)

Putrescine utilization regulator

Pyridoxal kinase (EC 2.7.1.35)

Pyridoxamine 5'-phosphate oxidase (EC 1.4.3.5)

Pyridoxine 5'-phosphate synthase (EC 2.6.99.2)

Pyridoxine biosynthesis glutamine amidotransferase, synthase subunit (EC 2.4.2.-)
Pyrimidine-nucleoside phosphorylase (EC 2.4.2.2)
Pyrophosphate-energized proton pump (EC 3.6.1.1)

Pyrophosphate-specific outer membrane porin OprO

Pyrroline-5-carboxylate reductase (EC 1.5.1.2)

Pyruvate carboxylase subunit B (biotin-containing) (EC 6.4.1.1)

Pyruvate dehydrogenase E1 component (EC 1.2.4.1)

Pyruvate dehydrogenase E1 component alpha subunit (EC 1.2.4.1)

Pyruvate dehydrogenase E1 component beta subunit (EC 1.2.4.1)
Pyruvate formate-lyase (EC 2.3.1.54)

Pyruvate formate-lyase activating enzyme (EC 1.97.1.4)
Pyruvate kinase (EC 2.7.1.40)

Pyruvate oxidase [ubiquinone, cytochrome] (EC 1.2.2.2)
Pyruvate-flavodoxin oxidoreductase (EC 1.2.7.-)
Pyruvate,phosphate dikinase (EC 2.7.9.1)
Queuosine Biosynthesis QueC ATPase
Queuosine biosynthesis QueD, PTPS-I

Queuosine Biosynthesis QueE Radical SAM

Quinate/shikimate 5-dehydrogenase I delta (EC 1.1.1.25)

Quinolinate phosphoribosyltransferase [decarboxylating] (EC 2.4.2.19)

Quinolinate synthetase (EC 2.5.1.72)

Phosphate:acyl-ACP acyltransferase PlsX

Phosphatidate cytidylyltransferase (EC 2.7.7.41)

Phosphatidylglycerophosphatase B (EC 3.1.3.27)

Phosphatidylserine decarboxylase (EC 4.1.1.65)

Phospho-N-acetylmuramoyl-pentapeptide-transferase (EC 2.7.8.13)

Phosphoadenylyl-sulfate reductase [thioredoxin] (EC 1.8.4.8)

Phosphoanhydride phosphohydrolase (EC 3.1.3.2) (pH 2.5 acid phosphatase) (AP)

Phosphoenolpyruvate carboxykinase [ATP] (EC 4.1.1.49)

Phosphoenolpyruvate carboxykinase [GTP] (EC 4.1.1.32)
Phosphoenolpyruvate carboxylase (EC 4.1.1.31)
Phosphoenolpyruvate synthase (EC 2.7.9.2)
Phosphoenolpyruvate-dihydroxyacetone phosphotransferase (EC 2.7.1.121), ADP-binding subunit DhaL

Phosphoenolpyruvate-protein phosphotransferase of PTS system (EC 2.7.3.9)

Phosphoethanolamine transferase EptA specific for the 1 phosphate group of core-lipid A

Phosphoethanolamine transferase specific for the outer Kdo residue of lipopolysaccharide

Phosphogluconate dehydratase (EC 4.2.1.12)

Phosphoglucosamine mutase (EC 5.4.2.10)
Phosphoglycerate kinase (EC 2.7.2.3)

Phosphoglycerate kinase (EC 2.7.2.3)
Phosphoglycerol transferase I (EC 2.7.8.20)

Phosphoglycolate phosphatase (EC 3.1.3.18)
Phosphoheptose isomerase (EC 5.3.1.-)
Phosphoheptose isomerase 1 (EC 5.3.1.-)
Phospholipid-lipopolysaccharide ABC transporter
Phosphomannomutase (EC 5.4.2.8)

Phosphonate ABC transporter ATP-binding protein (TC 3.A.1.9.1)

Phosphonate ABC transporter permease protein phnE (TC 3.A.1.9.1)

Phosphonate ABC transporter permease protein phnE1 (TC 3.A.1.9.1)

Phosphonate ABC transporter permease protein phnE2 (TC 3.A.1.9.1)

Radical SAM family enzyme, similar to coproporphyrinogen III oxidase, oxygen-independent, clustered with nucleoside-triphosphatase RdgB
Rare lipoprotein A precursor
RecA protein
RecD-like DNA helicase YrrC
Recombination inhibitory protein MutS2
Recombination protein RecR
Redox-sensitive transcriptional regulator (AT-rich DNA-binding protein)
Regulator of nucleoside diphosphate kinase
Regulatory protein of benzoate catabolism
Replicative DNA helicase (DnaB) (EC 3.6.4.12)
Replicative DNA helicase (EC 3.6.1.-) [SA14-24]
Respiratory nitrate reductase alpha chain (EC 1.7.99.4)
Respiratory nitrate reductase beta chain (EC 1.7.99.4)
Respiratory nitrate reductase delta chain (EC 1.7.99.4)
Respiratory nitrate reductase gamma chain (EC 1.7.99.4)
response regulator in two-component regulatory system with PhoQ
Response regulator NasT
Rhodanese domain protein UPF0176
Rhodanese domain protein UPF0176, cyanobacterial/alphaproteobacterial subgroup
Rhodanese domain protein, Enterobacterial subgroup, YceA homolog
Rhodanese-like domain protein
Rhodanese-like domain required for thiamine synthesis
Rhodanese-related sulfurtransferase
Riboflavin kinase (EC 2.7.1.26)
Riboflavin synthase eubacterial/eukaryotic (EC 2.5.1.9)
Ribokinase (EC 2.7.1.15)
Ribonuclease BN (EC 3.1.-.-)
Ribonuclease D (EC 3.1.26.3)
Ribonuclease E (EC 3.1.26.12)

Phosphonate ABC transporter phosphate-binding periplasmic component (TC 3.A.1.9.1)
Phosphonates transport ATP-binding protein PhnK
Phosphonates transport ATP-binding protein PhnL
Phosphopantetheine adenylyltransferase (EC 2.7.7.3)
Phosphopantothenoylcysteine decarboxylase (EC 4.1.1.36)
Phosphopantothenoylcysteine synthetase (EC 6.3.2.5)
Phosphopentomutase (EC 5.4.2.7)
Phosphoribosyl-AMP cyclohydrolase (EC 3.5.4.19)
Phosphoribosyl-ATP pyrophosphatase (EC 3.6.1.31)
Phosphoribosylamine--glycine ligase (EC 6.3.4.13)
Phosphoribosylaminoimidazole carboxylase ATPase subunit (EC 4.1.1.21)
Phosphoribosylaminoimidazole carboxylase catalytic subunit (EC 4.1.1.21)
Phosphoribosylaminoimidazole-succinocarboxamide synthase (EC 6.3.2.6)
Phosphoribosylaminoimidazolecarboxamide formyltransferase (EC 2.1.2.3)
Phosphoribosylanthranilate isomerase (EC 5.3.1.24)
Phosphoribosylformimino-5-aminoimidazole carboxamide ribotide isomerase (EC 5.3.1.16)
Phosphoribosylformylglycinamidine cyclo-ligase (EC 6.3.3.1)
Phosphoribosylformylglycinamidine synthase, glutamine amidotransferase subunit (EC 6.3.5.3)
Phosphoribosylformylglycinamidine synthase, synthetase subunit (EC 6.3.5.3)
Phosphoribosylglycinamide formyltransferase (EC 2.1.2.2)
Phosphoribosylglycinamide formyltransferase 2 (EC 2.1.2.-)
Phosphoserine aminotransferase (EC 2.6.1.52)
Phosphoserine phosphatase (EC 3.1.3.3)
Phytochrome, two-component sensor histidine kinase (EC 2.7.3.-)
Phytoene dehydrogenase (EC 1.14.99.-)
Phytoene dehydrogenase and related proteins
Phytoene desaturase (EC 1.14.99.-)
Phytoene synthase (EC 2.5.1.32)
Plasmid replication protein RepA

Ribonuclease E inhibitor RraA
Ribonuclease HI (EC 3.1.26.4)
Ribonuclease HII (EC 3.1.26.4)
Ribonuclease HIII (EC 3.1.26.4)
Ribonuclease I precursor (EC 3.1.27.6)
Ribonuclease III (EC 3.1.26.3)
Ribonuclease P protein component (EC 3.1.26.5)
Ribonuclease PH (EC 2.7.7.56)
Ribonuclease Z (EC 3.1.26.11)
Ribonucleotide reductase of class Ia (aerobic), alpha subunit (EC 1.17.4.1)
Ribonucleotide reductase of class Ia (aerobic), beta subunit (EC 1.17.4.1)
Ribonucleotide reductase of class Ib (aerobic), alpha subunit (EC 1.17.4.1)
Ribonucleotide reductase of class II (coenzyme B12-dependent) (EC 1.17.4.1)
Ribonucleotide reductase of class III (anaerobic), large subunit (EC 1.17.4.2)
Ribonucleotide reductase transcriptional regulator NrdR
Ribose 5-phosphate isomerase A (EC 5.3.1.6)
Ribose ABC transport system, ATP-binding protein RbsA (TC 3.A.1.2.1)
Ribose ABC transport system, high affinity permease RbsD (TC 3.A.1.2.1)
Ribose ABC transport system, periplasmic ribose-binding protein RbsB (TC 3.A.1.2.1)
Ribose ABC transport system, permease protein RbsC (TC 3.A.1.2.1)
Ribose ABC transporter, periplasmic ribose-binding protein RbsB (TC 3.A.1.2.1)
Ribose operon repressor
Ribose-phosphate pyrophosphokinase (EC 2.7.6.1)
Ribosomal large subunit pseudouridine synthase A (EC 4.2.1.70)
Ribosomal large subunit pseudouridine synthase B (EC 4.2.1.70)
Ribosomal large subunit pseudouridine synthase C (EC 4.2.1.70)
Ribosomal large subunit pseudouridine synthase D (EC 4.2.1.70)
Ribosomal large subunit pseudouridine synthase F (EC 4.2.1.70)
Ribosomal protein L11 methyltransferase (EC 2.1.1.-)

Plasmid replication protein RepB
Plasmid replication protein RepC
Poly(A) polymerase (EC 2.7.7.19)
Polyferredoxin NapH (periplasmic nitrate reductase)
polyhydroxyalkanoate synthesis repressor PhaR
Polyhydroxyalkanoic acid synthase
Polymyxin resistance protein ArnC, glycosyl transferase (EC 2.4.-.-)
Polymyxin resistance protein ArnT, undecaprenyl phosphate-alpha-L-Ara4N transferase
Polymyxin resistance protein PmrJ, predicted deacetylase
Polyphosphate kinase (EC 2.7.4.1)
Polyribonucleotide nucleotidyltransferase (EC 2.7.7.8)
Polysaccharide export lipoprotein Wza
Porphobilinogen synthase (EC 4.2.1.24)
Positive regulator of CheA protein activity (CheW)
Positive regulator of L-idonate catabolism
Positive regulator of Tartrate dehydrogenase/decarboxylase/D-malic enzyme
Positive transcription regulator EvgA
Possible hypoxanthine oxidase XdhD (EC 1.-.-.-)
Possible pyrimidine permease in reductive pathway
Possible pyrimidine-degrading protein DUF1688
Potassium channel protein
Potassium efflux system KefA protein
Potassium uptake protein TrkH
Potassium voltage-gated channel subfamily KQT
Potassium-transporting ATPase A chain (EC 3.6.3.12) (TC 3.A.3.7.1)
Potassium-transporting ATPase B chain (EC 3.6.3.12) (TC 3.A.3.7.1)
Potassium-transporting ATPase C chain (EC 3.6.3.12) (TC 3.A.3.7.1)
POTASSIUM/PROTON ANTIPORTER ROSB
Precorrin-6A synthase (EC 2.1.1.152)

ribosomal protein L7Ae family protein

Ribosomal protein S12p Asp88 (E. coli) methylthiotransferase

Ribosomal protein S6 glutaminyl transferase

Ribosomal RNA small subunit methyltransferase E (EC 2.1.1.-)

Ribosomal subunit interface protein

Ribosomal-protein-S18p-alanine acetyltransferase (EC 2.3.1.-)

Ribosomal-protein-S5p-alanine acetyltransferase

Ribosome modulation factor

Ribosome small subunit-stimulated GTPase EngC

Ribosome-associated heat shock protein implicated in the recycling of the 50S subunit (S4 paralog)

Ribosome-binding factor A

Ribulokinase (EC 2.7.1.16)

Ribulose bisphosphate carboxylase large chain (EC 4.1.1.39)

Ribulose-phosphate 3-epimerase (EC 5.1.3.1)

RNA binding protein, contains ribosomal protein S1 domain

RNA methyltransferase, TrmA family

RNA polymerase sigma factor

RNA polymerase sigma factor for flagellar operon

RNA polymerase sigma factor RpoD

RNA polymerase sigma factor RpoE

RNA polymerase sigma factor RpoH

RNA polymerase sigma factor RpoS

RNA polymerase sigma-54 factor RpoN

RNA polymerase sigma-70 factor

RNA polymerase sigma-H factor AlgT

RNA-binding protein Hfq

RNA:NAD 2'-phosphotransferase

RND efflux system, inner membrane transporter CmeB

Predicted 2-keto-4-pentenoate hydratase/2-oxohepta-3-ene-1,7-dioic acid hydratase

Predicted alpha-ribazole-5-phosphate synthase CblS for cobalamin biosynthesis

Predicted amidohydrolase RutB in novel pyrimidine catabolism pathway

Predicted ATP-dependent endonuclease of the OLD family, YbjD subgroup

Predicted ATPase related to phosphate starvation-inducible protein PhoH

Predicted beta-glucoside-regulated ABC transport system, permease component 1, COG1175

Predicted beta-glucoside-regulated ABC transport system, sugar binding component, COG1653

predicted biotin regulatory protein BioR (GntR family)

Predicted biotin repressor from TetR family

Predicted cell-wall-anchored protein SasA (LPXTG motif)

Predicted cobalt transporter CbtA

Predicted D-glucarate or D-galactorate regulator, GntR family

Predicted D-lactate dehydrogenase, Fe-S protein, FAD/FMN-containing

Predicted erythritol ABC transporter 1, permease component 1

Predicted erythritol ABC transporter 1, permease component 2

Predicted galacto-N-biose-/lacto-N-biose I ABC transporter, periplasmic substrate-binding protein

Predicted galacto-N-biose-/lacto-N-biose I ABC transporter, permease component 1

Predicted galacto-N-biose-/lacto-N-biose I ABC transporter, permease component 2

Predicted gluconate TRAP family transporter, DctM subunit

Predicted hydrolase of the metallo-beta-lactamase superfamily, clustered with KDO2-Lipid A biosynthesis genes

Predicted hydroxymethylpyrimidine transporter CytX

Predicted L-lactate dehydrogenase, Fe-S oxidoreductase subunit YkgE

Predicted L-lactate dehydrogenase, hypothetical protein subunit YkgG

Predicted L-lactate dehydrogenase, Iron-sulfur cluster-binding subunit YkgF

Predicted L-rhamnose ABC transporter, transmembrane component 1

Predicted Lactate-responsive regulator, IclR family

Predicted maltose transporter MalT

Predicted monooxygenase RutA in novel pyrimidine catabolism pathway

RND efflux system, membrane fusion protein CmeA
RND efflux system, outer membrane lipoprotein CmeC
RND efflux system, outer membrane lipoprotein, NodT family
Rod shape-determining protein MreB
Rod shape-determining protein MreC
Rod shape-determining protein RodA
Rossmann fold nucleotide-binding protein Smf possibly involved in DNA uptake
Rrf2 family transcriptional regulator
rRNA small subunit methyltransferase H
rRNA small subunit methyltransferase I
Rubisco activation protein CbbQ
RuBisCO operon transcriptional regulator CbbR
Rubrerythrin
S-(hydroxymethyl)glutathione dehydrogenase (EC 1.1.1.284)
S-adenosylmethionine decarboxylase proenzyme (EC 4.1.1.50), prokaryotic class 1A
S-adenosylmethionine synthetase (EC 2.5.1.6)
S-adenosylmethionine:tRNA ribosyltransferase-isomerase (EC 5.-.-.-)
S-formylglutathione hydrolase (EC 3.1.2.12)
salicylate esterase
Salicylate hydroxylase (EC 1.14.13.1)
SAM-dependent methyltransferase 2, in cluster with Hydroxyacylglutathione hydrolase (EC 3.1.2.6)
Sarcosine oxidase alpha subunit (EC 1.5.3.1)
Sarcosine oxidase beta subunit (EC 1.5.3.1)
Sarcosine oxidase gamma subunit (EC 1.5.3.1)
secreted alkaline phosphatase
Segregation and condensation protein A
Segregation and condensation protein B
Selenide,water dikinase (EC 2.7.9.3)

Predicted NAD regulator in Alphaproteobacteria
Predicted P-loop ATPase fused to an acetyltransferase COG1444
Predicted regulator PutR for proline utilization, GntR family
Predicted signal-transduction protein containing cAMP-binding and CBS domains
Predicted sugar ABC transport system, periplasmic binding protein YphF precursor
Predicted transcriptional regulator LiuR of leucine degradation pathway, MerR family
Predicted transcriptional regulator of N-Acetylglucosamine utilization, GntR family
Predicted transcriptional regulator of pyridoxine metabolism
Predicted transcriptional regulator of the myo-inositol catabolic operon
Predicted trehalose permease, MFS family
Prephenate dehydratase (EC 4.2.1.51)
Primosomal replication protein N
Probable 3-phenylpropionic acid transporter
Probable Co/Zn/Cd efflux system membrane fusion protein
probable dibenzothiophene desulfurization enzyme
Probable electron transfer flavoprotein-quinone oxidoreductase FixC (EC 1.5.5.-)
Probable glutathione S-transferase (EC 2.5.1.18), YfcF homolog
probable iron binding protein from the HesB_IscA_SufA family
Probable L-ascorbate-6-phosphate lactonase UlaG (EC 3.1.1.-) (L-ascorbate utilization protein G)
Probable low-affinity inorganic phosphate transporter
Probable monothiol glutaredoxin GrlA
Probable poly(beta-D-mannuronate) O-acetylase (EC 2.3.1.-)
Probable RND efflux membrane fusion protein
probable RuBisCo-expression protein CbbX
Probable transcription regulator protein of MDR efflux pump cluster
Probable VANILLIN dehydrogenase oxidoreductase protein (EC 1.-.-.-)
Proline dehydrogenase (EC 1.5.99.8) (Proline oxidase)
Proline iminopeptidase (EC 3.4.11.5)

Sensor histidine kinase colocalized with HrtAB transporter
Sensor protein basS/pmrB (EC 2.7.3.-)
Sensor protein basS/pmrB (EC 2.7.3.-)
Sensor protein PhoQ (EC 2.7.13.3)
Sensory histidine kinase AtoS
Sensory histidine kinase QseC
Septum formation protein Maf
Septum site-determining protein MinC
Septum site-determining protein MinD
Ser/Thr protein phosphatase family protein, UDP-2,3-diacylglucosamine hydrolase (EC 3.6.1.54) homolog
Serine acetyltransferase (EC 2.3.1.30)
Serine hydroxymethyltransferase (EC 2.1.2.1)
Serine protease precursor MucD/AlgY associated with sigma factor RpoE
Serine transporter
Serine--glyoxylate aminotransferase (EC 2.6.1.45)
Seryl-tRNA synthetase (EC 6.1.1.11)
Shikimate 5-dehydrogenase I alpha (EC 1.1.1.25)
Shikimate kinase I (EC 2.7.1.71)
Sialic acid transporter (permease) NanT
Sialic acid utilization regulator, RpiR family
Sialic acid-regulated TonB-dependent outer membrane receptor
Sialidase (EC 3.2.1.18)
Sigma factor RpoE negative regulatory protein RseB precursor
Signal peptidase I (EC 3.4.21.89)
Signal recognition particle receptor protein FtsY (=alpha subunit) (TC 3.A.5.1.1)
Signal recognition particle, subunit Ffh SRP54 (TC 3.A.5.1.1)
Signal transduction histidine kinase CheA (EC 2.7.3.-)
Signal transduction histidine kinase HoxJ (hydrogenase regulation)
Similar to phosphoglycolate phosphatase, clustered with ribosomal large subunit pseudouridine synthase C
Similar to phosphoglycolate phosphatase, clustered with ubiquinone biosynthesis

Proline-specific permease proY
Proline/sodium symporter PutP (TC 2.A.21.2.1)
Propionate catabolism operon regulatory protein PrpR
Propionate--CoA ligase (EC 6.2.1.17)
Propionyl-CoA carboxylase beta chain (EC 6.4.1.3)
Propionyl-CoA carboxylase carboxyl transferase subunit (EC 6.4.1.3)
Proposed lipoate regulatory protein YbeD
Protease VII (Omptin) precursor (EC 3.4.23.49)
Proteasome subunit beta (EC 3.4.25.1), bacterial
Protein acetyltransferase
Protein AraJ precursor
Protein export cytoplasm chaperone protein (SecB, maintains protein to be exported in unfolded state)
Protein involved in catabolism of external DNA
Protein of unknown function DUF1446
Protein of unknown function DUF374
Protein RcsF
Protein RtcB
Protein YhjJ, putative peptidase
Protein YigP (COG3165) clustered with ubiquinone biosynthetic genes
Protein yihD
Protein-L-isoaspartate O-methyltransferase (EC 2.1.1.77)
Protein-N(5)-glutamine methyltransferase PrmB, methylates LSU ribosomal protein L3p
Protein-N(5)-glutamine methyltransferase PrmC, methylates polypeptide chain release factors RF1 and RF2
Protocatechuate 3,4-dioxygenase beta chain (EC 1.13.11.3)
Protocatechuate 4,5-dioxygenase alpha chain (EC 1.13.11.8)
Protocatechuate 4,5-dioxygenase beta chain (EC 1.13.11.8)
Pseudouridine 5'-phosphate glycosidase
Pterin-4-alpha-carbinolamine dehydratase (EC 4.2.1.96)
PTS system, 2-O-alpha-mannosyl-D-glycerate-specific IIB component
PTS system, arbutin-, cellobiose-, and salicin-specific IIBC component (EC 2.7.1.69)

SAM-dependent O-methyltransferase
Similar to TadZ/CpaE, associated with Flp pilus assembly
Similarity with glutathionylspermidine synthase (EC 6.3.1.8), group 1
Single-stranded DNA-binding protein
Single-stranded-DNA-specific exonuclease RecJ (EC 3.1.-.-)
Site-specific tyrosine recombinase
Sodium-Choline Symporter
Sodium-dependent phosphate transporter
Soluble lytic murein transglycosylase precursor (EC 3.2.1.-)
Soluble pyridine nucleotide transhydrogenase (EC 1.6.1.1)
Sorbitol operon transcription regulator
SOS-response repressor and protease LexA (EC 3.4.21.88)
Spermidine N1-acetyltransferase (EC 2.3.1.57)
Spermidine Putrescine ABC transporter permease component PotB (TC 3.A.1.11.1)
Spermidine Putrescine ABC transporter permease component potC (TC_3.A.1.11.1)
Spermidine synthase (EC 2.5.1.16)
Spore germination protein GerKB
SSU ribosomal protein S10p (S20e)
SSU ribosomal protein S11p (S14e)
SSU ribosomal protein S12p (S23e)
SSU ribosomal protein S12p (S23e)
SSU ribosomal protein S13p (S18e)
SSU ribosomal protein S14p (S29e)
SSU ribosomal protein S15p (S13e)
SSU ribosomal protein S16p
SSU ribosomal protein S17p (S11e)
SSU ribosomal protein S19p (S15e)
SSU ribosomal protein S1p
SSU ribosomal protein S20p

PTS system, beta-glucoside-specific IIB component (EC 2.7.1.69)
PTS system, cellobiose-specific IIC component (EC 2.7.1.69)
PTS system, fructose-specific IIB component (EC 2.7.1.69)
PTS system, fructose-specific IIB component (EC 2.7.1.69)
PTS system, galactitol-specific IIA component (EC 2.7.1.69)
PTS system, galactitol-specific IIC component (EC 2.7.1.69)
PTS system, glucitol/sorbitol-specific IIB component and second of two IIC components (EC 2.7.1.69)
PTS system, maltose and glucose-specific IIC component (EC 2.7.1.69)
PTS system, mannitol-specific IIB component (EC 2.7.1.69)
PTS system, mannose-specific IIB component (EC 2.7.1.69)
PTS system, mannose-specific IIB component (EC 2.7.1.69)
PTS system, N-acetylglucosamine-specific IIB component (EC 2.7.1.69)
PTS system, N-acetylmuramic acid-specific IIB component (EC 2.7.1.69)
PTS system, N,N'-diacetylchitobiose-specific IIC component
Purine nucleoside phosphorylase (EC 2.4.2.1)
Purine nucleotide synthesis repressor
Putative 2-keto-3-deoxygluconate kinase (EC 2.7.1.45)
Putative capsular polysaccharide transport protein YegH
putative Cytochrome bd2, subunit I
putative Cytochrome bd2, subunit II
Putative cytochrome C-type biogenesis protein
putative cytochrome P450 hydroxylase
Putative deoxyribonuclease YcfH
Putative deoxyribose-specific ABC transporter, ATP-binding protein
Putative diaminopropionate ammonia-lyase (EC 4.3.1.15)
Putative diheme cytochrome c-553
Putative Dihydrolipoamide dehydrogenase (EC 1.8.1.4)
Putative dihydroxyacetone kinase (EC 2.7.1.29), dihydroxyacetone binding subunit

SSU ribosomal protein S21p
SSU ribosomal protein S2p (SAe)
SSU ribosomal protein S3p (S3e)
SSU ribosomal protein S4p (S9e)
SSU ribosomal protein S5p (S2e)
SSU ribosomal protein S6p
SSU ribosomal protein S7p (S5e)
SSU ribosomal protein S8p (S15Ae)
SSU ribosomal protein S9p (S16e)
Stage V sporulation protein involved in spore cortex synthesis (SpoVR)
Starvation sensing protein RspA
Stringent starvation protein A
Substrate-specific component BioY of biotin ECF transporter
Succinate dehydrogenase cytochrome b subunit
Succinate dehydrogenase flavoprotein subunit (EC 1.3.99.1)
Succinate dehydrogenase hydrophobic membrane anchor protein
Succinate dehydrogenase iron-sulfur protein (EC 1.3.99.1)
Succinate-semialdehyde dehydrogenase [NADP+] (EC 1.2.1.79)
Succinyl-CoA ligase [ADP-forming] alpha chain (EC 6.2.1.5)
Succinyl-CoA ligase [ADP-forming] beta chain (EC 6.2.1.5)
Succinyl-CoA:3-ketoacid-coenzyme A transferase subunit A (EC 2.8.3.5)
Succinyl-CoA:3-ketoacid-coenzyme A transferase subunit B (EC 2.8.3.5)
Succinylarginine dihydrolase (EC 3.5.3.23)
Succinylglutamate desuccinylase (EC 3.5.1.96)
Succinylglutamic semialdehyde dehydrogenase (EC 1.2.1.71)
Succinylornithine transaminase (EC 2.6.1.81)
Sucrose operon repressor ScrR, LacI family
Sugar phosphate isomerases/epimerases family protein YcjR
Sugar/maltose fermentation stimulation protein homolog

putative esterase
Putative formate dehydrogenase oxidoreductase protein
Putative formate dehydrogenase-specific chaperone
Putative fumarylacetoacetate (FAA) hydrolase
Putative glutathione transporter, ATP-binding component
Putative glutathione transporter, permease component
Putative glutathione transporter,solute-binding component
putative Glutathione-regulated potassium-efflux system protein KefB
Putative gluthatione transporter,solute-binding component
Putative hemin-binding lipoprotein
Putative inner membrane protein YjeT (clustered with HflC)
Putative iron-sulfur cluster assembly scaffold protein for SUF system, SufE2
Putative lipase in cluster with Phosphatidate cytidylyltransferase
Putative membrane protein, clustering with ActP
Putative membrane transporter ATPase,YhiD
Putative metal chaperone, involved in Zn homeostasis, GTPase of COG0523 family
Putative n-hydroxybenzoate hydroxylase
Putative n-hydroxybenzoate hydroxylase
Putative outer membrane TonB-dependent receptor associated with haemagglutinin family outer membrane protein
Putative oxidoreductase YcjS (EC 1.-.-.-), NADH-binding
Putative oxidoreductase YdjL
Putative oxidoreductase YncB
Putative phosphoenolpyruvate synthase/pyruvate phosphate dikinase, C-terminal domain
Putative phosphoenolpyruvate synthase/pyruvate phosphate dikinase, N-terminal domain
Putative predicted metal-dependent hydrolase
Putative preQ0 transporter
Putative sensor-like histidine kinase YfhK
putative sodium-dependent bicarbonate transporter
Putative stomatin/prohibitin-family membrane protease subunit YbbK

Sulfate adenylyltransferase subunit 1 (EC 2.7.7.4)

Sulfate adenylyltransferase subunit 2 (EC 2.7.7.4)

Sulfate and thiosulfate binding protein CysP

Sulfate and thiosulfate import ATP-binding protein CysA (EC 3.6.3.25)

Sulfate permease

Sulfate transport system permease protein CysT

Sulfate transport system permease protein CysW

Sulfate-binding protein Sbp

Sulfide dehydrogenase [flavocytochrome C] flavoprotein chain precursor (EC 1.8.2.-)

Sulfite reductase [NADPH] hemoprotein beta-component (EC 1.8.1.2)

Sulfur carrier protein adenylyltransferase ThiF

Sulfur oxidation molybdopterin C protein

Sulfur oxidation protein SoxB

Sulfur oxidation protein SoxX

Sulfur oxidation protein SoxZ

Superoxide dismutase [Cu-Zn] precursor (EC 1.15.1.1)

Superoxide dismutase [Fe] (EC 1.15.1.1)

Survival protein SurA precursor (Peptidyl-prolyl cis-trans isomerase SurA) (EC 5.2.1.8)

Tagatose 1,6-bisphosphate aldolase (EC 4.1.2.40)

Tagatose-6-phosphate kinase (EC 2.7.1.144)

Tagatose-6-phosphate kinase AgaZ (EC 2.7.1.144)

Taurine transport system permease protein TauC

TcuA: flavoprotein used to oxidize tricarballylate to cis-aconitate

TcuB: works with TcuA to oxidize tricarballylate to cis-aconitate

TcuC: integral membrane protein used to transport tricarballylate across the cell membrane

Teichoic acid export ATP-binding protein TagH (EC 3.6.3.40)

Teicoplanin-resistance associated HTH-type transcriptional regulator TcaR

Tetraacyldisaccharide 4'-kinase (EC 2.7.1.130)

Thiamin ABC transporter, transmembrane component

Putative sucrose phosphorylase (EC 2.4.1.7)

Putative two-component system response regulator YedW

Putative two-domain glycosyltransferase

PutR, transcriptional activator of PutA and PutP

Putrescine ABC transporter putrescine-binding protein PotF (TC 3.A.1.11.2)

Putrescine aminotransferase (EC 2.6.1.82)

Putrescine importer

Putrescine transport ATP-binding protein PotA (TC 3.A.1.11.1)

Putrescine transport system permease protein PotH (TC 3.A.1.11.2)

Putrescine utilization regulator

Putrescine/proton symporter, putrescine/ornithine antiporter PotE

Pyridine nucleotide-disulphide oxidoreductase associated with reductive pyrimidine catabolism

Pyridoxal kinase (EC 2.7.1.35)

Pyridoxamine 5'-phosphate oxidase (EC 1.4.3.5)

Pyridoxine 5'-phosphate synthase (EC 2.6.99.2)

Pyridoxine biosynthesis glutamine amidotransferase, glutaminase subunit (EC 2.4.2.-)

Pyridoxine biosynthesis glutamine amidotransferase, synthase subunit (EC 2.4.2.-)

Pyrophosphate-dependent fructose 6-phosphate-1-kinase (EC 2.7.1.90)

Pyrophosphate-energized proton pump (EC 3.6.1.1)

Pyrroline-5-carboxylate reductase (EC 1.5.1.2)

Pyruvate decarboxylase (EC 4.1.1.1)

Pyruvate dehydrogenase E1 component (EC 1.2.4.1)

Pyruvate dehydrogenase E1 component alpha subunit (EC 1.2.4.1)

Pyruvate dehydrogenase E1 component beta subunit (EC 1.2.4.1)

Pyruvate formate-lyase (EC 2.3.1.54)

Pyruvate formate-lyase activating enzyme (EC 1.97.1.4)

Pyruvate kinase (EC 2.7.1.40)

Pyruvate oxidase [ubiquinone, cytochrome] (EC 1.2.2.2)

Pyruvate-flavodoxin oxidoreductase (EC 1.2.7.-)

Thiamin-phosphate pyrophosphorylase (EC 2.5.1.3)
Thiaminase II (EC 3.5.99.2) involved in salvage of thiamin pyrimidine moiety
Thiamine biosynthesis protein thiI
Thiamine-monophosphate kinase (EC 2.7.4.16)
Thiazole biosynthesis protein ThiG
Thiazole tautomerase TenI (EC 5.3.99.10)
ThiJ/PfpI family protein
Thiol peroxidase, Bcp-type (EC 1.11.1.15)
Thiol:disulfide interchange protein DsbC
Thioredoxin
Thioredoxin reductase (EC 1.8.1.9)
Thiosulfate sulfurtransferase, rhodanese (EC 2.8.1.1)
Threonine dehydratase (EC 4.3.1.19)
Threonine dehydratase biosynthetic (EC 4.3.1.19)
Threonine dehydratase, catabolic (EC 4.3.1.19)
Threonine dehydrogenase and related Zn-dependent dehydrogenases
Threonine synthase (EC 4.2.3.1)
Threonyl-tRNA synthetase (EC 6.1.1.3)
Thymidine phosphorylase (EC 2.4.2.4)
Thymidylate kinase (EC 2.7.4.9)
Thymidylate synthase (EC 2.1.1.45)
TldE/PmbA family protein, Beta/Gamma-proteobacterial subgroup
tmRNA-binding protein SmpB
TolA protein
tolB protein precursor, periplasmic protein involved in the tonb-independent uptake of group A colicins
Toluene-4-monooxygenase, subunit TmoA
toluenesulfonate zinc-independent alcohol dehydrogenase
TonB-dependent receptor
TonB-dependent siderophore receptor
Topoisomerase IV subunit A (EC 5.99.1.-)

Pyruvate,phosphate dikinase (EC 2.7.9.1)
Queuosine Biosynthesis QueC ATPase
Queuosine biosynthesis QueD, PTPS-I
Queuosine Biosynthesis QueE Radical SAM
quinol oxidase polypeptide II QoxA (EC:1.9.3.-)
Quinolinate phosphoribosyltransferase [decarboxylating] (EC 2.4.2.19)
Quinolinate synthetase (EC 2.5.1.72)
Quinone oxidoreductase (EC 1.6.5.5)
quinoprotein alcohol dehydrogenase
Radical SAM domain heme biosynthesis protein
Radical SAM family enzyme, similar to coproporphyrinogen III oxidase, oxygen-independent, clustered with nucleoside-triphosphatase RdgB
Rare lipoprotein A precursor
RecA protein
RecD-like DNA helicase YrrC
Recombination inhibitory protein MutS2
Recombination protein RecR
Redox-sensitive transcriptional regulator (AT-rich DNA-binding protein)
regulator of length of O-antigen component of lipopolysaccharide chains
Regulator of nucleoside diphosphate kinase
Regulatory protein (induces abgABT, used to catabolize p-aminobenzoyl-glutamate)
Regulatory protein of benzoate catabolism
Replicative DNA helicase (DnaB) (EC 3.6.4.12)
Replicative DNA helicase (EC 3.6.1.-) [SA14-24]
Respiratory nitrate reductase alpha chain (EC 1.7.99.4)
Respiratory nitrate reductase beta chain (EC 1.7.99.4)
Respiratory nitrate reductase gamma chain (EC 1.7.99.4)
Response regulator BaeR
response regulator in two-component regulatory system with PhoQ
Response regulator of zinc sigma-54-dependent two-component system
Retron-type reverse transcriptase

Topoisomerase IV subunit B (EC 5.99.1.-)
TPR repeat containing exported protein
Transaldolase (EC 2.2.1.2)
Transcription antitermination protein NusG
Transcription elongation factor GreA
Transcription elongation factor GreB
Transcription repressor of multidrug efflux pump acrAB operon, TetR (AcrR) family
Transcription termination factor Rho
Transcription termination protein NusA
Transcription termination protein NusB
Transcription-repair coupling factor
Transcriptional activator MetR
Transcriptional activator of acetoin dehydrogenase operon AcoR
Transcriptional activator of maltose regulon, MalT
Transcriptional regulator BkdR of isoleucine and valine catabolism operon
Transcriptional regulator GabR of GABA utilization (GntR family with aminotransferase-like domain)
Transcriptional regulator NanR
Transcriptional regulator of mannitol utilization, DeoR family protein
Transcriptional regulator PhnF
transcriptional regulator, Crp/Fnr family
Transcriptional regulator, IclR family
Transcriptional regulator, MerR family
Transcriptional regulatory protein RstA
Transcriptional repressor of aga operon
Transcriptional repressor protein TrpR
Transketolase (EC 2.2.1.1)
Transketolase, C-terminal section (EC 2.2.1.1)
Transketolase, N-terminal section (EC 2.2.1.1)
Translation elongation factor G
Translation elongation factor G-related protein
Translation elongation factor LepA

Retron-type reverse transcriptase
Rhamnogalacturonides degradation protein RhiN
Rhamnulokinase (EC 2.7.1.5)
Rhamnulose-1-phosphate aldolase (EC 4.1.2.19)
Rhodanese domain protein UPF0176, cyanobacterial/alphaproteobacterial subgroup
Rhodanese domain protein, Enterobacterial subgroup, YceA homolog
Rhodanese-like domain protein
Rhodanese-related sulfurtransferase
Riboflavin kinase (EC 2.7.1.26)
Riboflavin synthase eubacterial/eukaryotic (EC 2.5.1.9)
Riboflavin transporter PnuX
Ribokinase (EC 2.7.1.15)
Ribonuclease BN (EC 3.1.-.-)
Ribonuclease D (EC 3.1.26.3)
Ribonuclease E (EC 3.1.26.12)
Ribonuclease E inhibitor RraA
Ribonuclease E inhibitor RraB
Ribonuclease HI (EC 3.1.26.4)
Ribonuclease HII (EC 3.1.26.4)
Ribonuclease I precursor (EC 3.1.27.6)
Ribonuclease III (EC 3.1.26.3)
Ribonuclease P protein component (EC 3.1.26.5)
Ribonuclease PH (EC 2.7.7.56)
Ribonuclease PH (EC 2.7.7.56)
Ribonuclease T (EC 3.1.13.-)
Ribonuclease Z (EC 3.1.26.11)
Ribonucleotide reductase of class Ia (aerobic), alpha subunit (EC 1.17.4.1)
Ribonucleotide reductase of class Ia (aerobic), beta subunit (EC 1.17.4.1)
Ribonucleotide reductase of class Ib (aerobic), alpha subunit (EC 1.17.4.1)
Ribonucleotide reductase of class Ib (aerobic), beta subunit (EC 1.17.4.1)
Ribonucleotide reductase of class II (coenzyme B12-dependent) (EC 1.17.4.1)

Translation elongation factor P
Translation elongation factor Ts
Translation elongation factor Tu
Translation initiation factor 1
Translation initiation factor 2
Translation initiation factor 3
Translation initiation factor SUI1-related protein
Transmembrane component CbrV of energizing module of predicted cobalamin ECF transporter
Transmembrane component of general energizing module of ECF transporters
Transmembrane component YkoC of energizing module of thiamin-regulated ECF transporter for HydroxyMethylPyrimidine
Transport ATP-binding protein CydC
Transport ATP-binding protein CydCD
Transport ATP-binding protein CydD
TRAP transporter solute receptor, unknown substrate 1
TRAP transporter, 4TM/12TM fusion protein, unknown substrate 1
TRAP-type transport system, periplasmic component, predicted N-acetylneuraminate-binding protein
Trehalose operon transcriptional repressor
Trehalose synthase (EC 5.4.99.16)
Trehalose-6-phosphate phosphatase (EC 3.1.3.12)
Triacylglycerol lipase (EC 3.1.1.3)
Tricarboxylate transport membrane protein TctA
Trimethylamine-N-oxide reductase (EC 1.6.6.9)
Triosephosphate isomerase (EC 5.3.1.1)
Trk system potassium uptake protein TrkA
TrkA-N:Sodium/hydrogen exchanger
tRNA (5-methylaminomethyl-2-thiouridylate)-methyltransferase (EC 2.1.1.61)
tRNA (Guanine37-N1) -methyltransferase (EC 2.1.1.31)
tRNA (uracil(54)-C5)-methyltransferase (EC 2.1.1.35)

Ribonucleotide reductase of class III (anaerobic), large subunit (EC 1.17.4.2)
Ribonucleotide reductase transcriptional regulator NrdR
Ribonucleotide reduction protein NrdI
Ribose 5-phosphate isomerase A (EC 5.3.1.6)
Ribose 5-phosphate isomerase B (EC 5.3.1.6)
Ribose ABC transport system, ATP-binding protein RbsA (TC 3.A.1.2.1)
Ribose ABC transport system, high affinity permease RbsD (TC 3.A.1.2.1)
Ribose ABC transport system, periplasmic ribose-binding protein RbsB (TC 3.A.1.2.1)
Ribose ABC transport system, permease protein RbsC (TC 3.A.1.2.1)
Ribose ABC transporter, periplasmic ribose-binding protein RbsB (TC 3.A.1.2.1)
Ribose operon repressor
Ribose-phosphate pyrophosphokinase (EC 2.7.6.1)
Ribosomal large subunit pseudouridine synthase A (EC 4.2.1.70)
Ribosomal large subunit pseudouridine synthase B (EC 4.2.1.70)
Ribosomal large subunit pseudouridine synthase C (EC 4.2.1.70)
Ribosomal large subunit pseudouridine synthase D (EC 4.2.1.70)
Ribosomal large subunit pseudouridine synthase F (EC 4.2.1.70)
Ribosomal protein L11 methyltransferase (EC 2.1.1.-)
ribosomal protein L7Ae family protein
Ribosomal protein S12p Asp88 (E. coli) methylthiotransferase
Ribosomal protein S6 glutaminyl transferase
Ribosomal RNA large subunit methyltransferase K
Ribosomal RNA small subunit methyltransferase E (EC 2.1.1.-)
Ribosomal small subunit pseudouridine synthase A (EC 4.2.1.70)
Ribosomal subunit interface protein
Ribosomal-protein-S18p-alanine acetyltransferase (EC 2.3.1.-)
Ribosomal-protein-S5p-alanine acetyltransferase
Ribosome hibernation protein YhbH

tRNA nucleotidyltransferase (EC 2.7.7.21) (EC 2.7.7.25)
tRNA pseudouridine 13 synthase (EC 4.2.1.-)
tRNA pseudouridine synthase A (EC 4.2.1.70)
tRNA pseudouridine synthase B (EC 4.2.1.70)
tRNA S(4)U 4-thiouridine synthase (former ThiI)
tRNA-dependent lipid II--amino acid ligase
tRNA-dihydrouridine synthase C (EC 1.-.-.-)
tRNA-guanine transglycosylase (EC 2.4.2.29)
tRNA-i(6)A37 methylthiotransferase
tRNA-specific adenosine-34 deaminase (EC 3.5.4.-)
tRNA:Cm32/Um32 methyltransferase
tRNA(Ile)-lysidine synthetase (EC 6.3.4.19)
Tryptophan 2,3-dioxygenase (EC 1.13.11.11)
Tryptophan synthase alpha chain (EC 4.2.1.20)
Tryptophan synthase beta chain (EC 4.2.1.20)
Tryptophanase (EC 4.1.99.1)
Tryptophanyl-tRNA synthetase (EC 6.1.1.2)
TsaB protein, required for threonylcarbamoyladenosine (t(6)A) formation in tRNA
TsaC protein (YrdC domain) required for threonylcarbamoyladenosine t(6)A37 modification in tRNA
TsaD/Kae1/Qri7 protein, required for threonylcarbamoyladenosine t(6)A37 formation in tRNA
TsaE protein, required for threonylcarbamoyladenosine t(6)A37 formation in tRNA
tungsten-containing formate dehydrogenase beta subunit
Twin-arginine translocation protein TatC
Twitching motility protein PilT
Two-component nitrogen fixation transcriptional regulator FixJ
Two-component oxygen-sensor histidine kinase FixL
Two-component response regulator CreC
Two-component response regulator YvcP
Two-component sensor histidine kinase BceS

Ribosome small subunit-stimulated GTPase EngC
Ribosome-associated heat shock protein implicated in the recycling of the 50S subunit (S4 paralog)
Ribosome-binding factor A
Ribosyl nicotinamide transporter, PnuC-like
Ribulokinase (EC 2.7.1.16)
Ribulose bisphosphate carboxylase large chain (EC 4.1.1.39)
Ribulose bisphosphate carboxylase small chain (EC 4.1.1.39)
Ribulose-phosphate 3-epimerase (EC 5.1.3.1)
RNA 3'-terminal phosphate cyclase (EC 6.5.1.4)
RNA methyltransferase, TrmA family
RNA polymerase sigma factor
RNA polymerase sigma factor for flagellar operon
RNA polymerase sigma factor RpoD
RNA polymerase sigma factor RpoE
RNA polymerase sigma factor RpoH
RNA polymerase sigma factor RpoS
RNA polymerase sigma factor SigB
RNA polymerase sigma-54 factor RpoN
RNA polymerase sigma-70 factor
RNA-binding protein Hfq
RND efflux system, inner membrane transporter CmeB
RND efflux system, membrane fusion protein CmeA
RND efflux system, outer membrane lipoprotein CmeC
RND efflux system, outer membrane lipoprotein, NodT family
Rod shape-determining protein MreB
Rod shape-determining protein MreC
Rod shape-determining protein MreD
Rod shape-determining protein RodA
Rossmann fold nucleotide-binding protein Smf possibly involved in DNA uptake

Two-component sensor histidine kinase PleC

Two-component sensor kinase SA14-24

Two-component sensor PilS

Two-component system response regulator

Two-component system response regulator QseB

type 1 fimbriae anchoring protein FimD

type 1 fimbriae major subunit FimA

Type cbb3 cytochrome oxidase biogenesis protein CcoG, involved in Cu oxidation

Type I restriction-modification system, DNA-methyltransferase subunit M (EC 2.1.1.72)

Type I restriction-modification system, restriction subunit R (EC 3.1.21.3)

Type I secretion outer membrane protein, TolC precursor

Type II secretory pathway, ATPase PulE/Tfp pilus assembly pathway, ATPase PilB

Type II/IV secretion system ATP hydrolase TadA/VirB11/CpaF, TadA subfamily

Type II/IV secretion system ATPase TadZ/CpaE, associated with Flp pilus assembly

Type II/IV secretion system protein TadC, associated with Flp pilus assembly

Type II/IV secretion system secretin RcpA/CpaC, associated with Flp pilus assembly

Type III restriction-modification system methylation subunit (EC 2.1.1.72)

Type IV fimbriae expression regulatory protein PilR

Type IV fimbrial assembly protein PilC

Type IV fimbrial assembly, ATPase PilB

type IV pili signal transduction protein PilI

Type IV pilin PilA

Type IV pilus biogenesis protein PilM

Type IV pilus biogenesis protein PilN

Type IV pilus biogenesis protein PilP

Type IV pilus biogenesis protein PilQ

Type IV secretion system protein VirD4

Tyrosine recombinase XerC

Tyrosine-protein kinase Wzc (EC 2.7.10.2)

Rrf2 family transcriptional regulator, group III

rRNA small subunit 7-methylguanosine (m7G) methyltransferase GidB

rRNA small subunit methyltransferase H

rRNA small subunit methyltransferase I

Rubisco activation protein CbbQ

RuBisCO operon transcriptional regulator CbbR

Rubredoxin

Rubredoxin-NAD(+) reductase (EC 1.18.1.1)

Rubrerythrin

S-(hydroxymethyl)glutathione dehydrogenase (EC 1.1.1.284)

S-adenosylmethionine decarboxylase proenzyme (EC 4.1.1.50), prokaryotic class 1A

S-adenosylmethionine synthetase (EC 2.5.1.6)

S-adenosylmethionine-dependent methyltransferase Functionally Coupled to the MukBEF Chromosome Partitioning Mechanism

S-adenosylmethionine:tRNA ribosyltransferase-isomerase (EC 5.-.-.-)

S-formylglutathione hydrolase (EC 3.1.2.12)

S-ribosylhomocysteine lyase (EC 4.4.1.21)

salicylate esterase

Salicylate hydroxylase (EC 1.14.13.1)

SAM-dependent methyltransferase 2, in cluster with Hydroxyacylglutathione hydrolase (EC 3.1.2.6)

Sarcosine oxidase alpha subunit (EC 1.5.3.1)

Sarcosine oxidase beta subunit (EC 1.5.3.1)

secreted alkaline phosphatase

Segregation and condensation protein A

Segregation and condensation protein B

selenocysteine-containing

Selenocysteine-specific translation elongation factor

Selenophosphate-dependent tRNA 2-selenouridine synthase

Sensor histidine protein kinase UhpB, glucose-6-phosphate specific (EC 2.7.13.3)

Sensor kinase CitA, DpiB (EC 2.7.3.-)

Tyrosyl-tRNA synthetase (EC 6.1.1.1)
ubiquinol cytochrome C oxidoreductase, cytochrome C1 subunit
Ubiquinol--cytochrome c reductase, cytochrome B subunit (EC 1.10.2.2)
Ubiquinol-cytochrome C reductase iron-sulfur subunit (EC 1.10.2.2)
Ubiquinone biosynthesis monooxygenase UbiB
Ubiquinone/menaquinone biosynthesis methyltransferase UbiE (EC 2.1.1.-)
UDP-3-O-[3-hydroxymyristoyl] N-acetylglucosamine deacetylase (EC 3.5.1.108)
UDP-4-amino-4-deoxy-L-arabinose formyltransferase (EC 2.1.2.13)
UDP-4-amino-4-deoxy-L-arabinose--oxoglutarate aminotransferase (EC 2.6.1.-)
UDP-glucose 4-epimerase (EC 5.1.3.2)
UDP-glucose dehydrogenase (EC 1.1.1.22)
UDP-glucuronic acid oxidase (UDP-4-keto-hexauronic acid decarboxylating) (EC 1.1.1.305)
UDP-N-acetylenolpyruvoylglucosamine reductase (EC 1.1.1.158)
UDP-N-acetylglucosamine 1-carboxyvinyltransferase (EC 2.5.1.7)
UDP-N-acetylglucosamine 2-epimerase (EC 5.1.3.14)
UDP-N-acetylglucosamine 4-epimerase (EC 5.1.3.7)
UDP-N-acetylglucosamine 4,6-dehydratase (EC 4.2.1.-)
UDP-N-acetylglucosamine--N-acetylmuramyl-(pentapeptide) pyrophosphoryl-undecaprenol N-acetylglucosamine transferase (EC 2.4.1.227)
UDP-N-acetylmuramate--alanine ligase (EC 6.3.2.8)
UDP-N-acetylmuramoylalanine--D-glutamate ligase (EC 6.3.2.9)
UDP-N-acetylmuramoylalanyl-D-glutamate--2,6-diaminopimelate ligase (EC 6.3.2.13)
UDP-N-acetylmuramoylalanyl-D-glutamyl-2,6-diaminopimelate--D-alanyl-D-alanine ligase (EC 6.3.2.10)
Uncharacterized ABC transporter, periplasmic component YrbD
Uncharacterized glutathione S-transferase-like protein
Uncharacterized GST-like protein yncG
Uncharacterized monothiol glutaredoxin ycf64-like
Uncharacterized protein ImpJ/VasE
Sensor protein basS/pmrB (EC 2.7.3.-)
Sensor protein PhoQ (EC 2.7.13.3)
Sensor protein torS (EC 2.7.3.-)
Sensory histidine kinase BaeS
Sensory histidine kinase in two-component regulatory system with RstA
Sensory histidine kinase QseC
Septum formation protein Maf
Septum site-determining protein MinC
Septum site-determining protein MinD
Ser/Thr protein phosphatase family protein, UDP-2,3-diacylglucosamine hydrolase (EC 3.6.1.54) homolog
Serine acetyltransferase (EC 2.3.1.30)
Serine hydroxymethyltransferase (EC 2.1.2.1)
Serine protease precursor MucD/AlgY associated with sigma factor RpoE
Serine transporter
Serine--glyoxylate aminotransferase (EC 2.6.1.45)
Seryl-tRNA synthetase (EC 6.1.1.11)
SgrR, sugar-phosphate stress, transcriptional activator of SgrS small RNA
Shikimate 5-dehydrogenase I alpha (EC 1.1.1.25)
Shikimate kinase I (EC 2.7.1.71)
Shikimate kinase III (EC 2.7.1.71)
Short chain fatty acids transporter
Sialic acid transporter (permease) NanT
Sialic acid utilization regulator, RpiR family
Sialic acid-induced transmembrane protein YjhT(NanM), possible mutarotase
Sialic acid-regulated TonB-dependent outer membrane receptor
Sialidase (EC 3.2.1.18)
Sigma factor RpoE negative regulatory protein RseB precursor

Uncharacterized protein in putrescine utilization cluster

Uncharacterized protein YphG, TPR-domain containing

Uncharacterized protein, Bsl7517 homolog

Uncharacterized protein, similar to the N-terminal domain of Lon protease

Uncharacterized siderophore biosynthesis protein near heme transporter HtsABC

Undecaprenyl-phosphate galactosephosphotransferase (EC 2.7.8.6)

Undecaprenyl-phosphate N-acetylglucosaminyl 1-phosphate transferase (EC 2.7.8.-)

Universal stress protein F

Universal stress protein family

Universal stress protein family COG0589

UPF0234 protein YajQ

UPF0265 protein YeeX

Uptake hydrogenase large subunit (EC 1.12.99.6)

Uptake hydrogenase small subunit precursor (EC 1.12.99.6)

Uracil permease

Uracil phosphoribosyltransferase (EC 2.4.2.9)

Uracil-DNA glycosylase, family 1

Uracil-DNA glycosylase, putative family 6

Urea ABC transporter, ATPase protein UrtD

Urea ABC transporter, ATPase protein UrtE

Urea ABC transporter, permease protein UrtB

Urea ABC transporter, permease protein UrtC

Urea ABC transporter, urea binding protein

Urea carboxylase (EC 6.3.4.6)

Urea carboxylase-related ABC transporter, periplasmic substrate-binding protein

Urea carboxylase-related ABC transporter, permease protein

Urea carboxylase-related aminomethyltransferase (EC 2.1.2.10)

Urease accessory protein UreD

Urease accessory protein UreE

Sigma factor RpoE regulatory protein RseC

Signal peptidase I (EC 3.4.21.89)

Signal recognition particle receptor protein FtsY (=alpha subunit) (TC 3.A.5.1.1)

Signal recognition particle, subunit Ffh SRP54 (TC 3.A.5.1.1)

Signal transduction histidine kinase CheA (EC 2.7.3.-)

Similar to C-terminal Zn-finger domain of DNA topoisomerase I

Similar to phosphoglycolate phosphatase, clustered with ribosomal large subunit pseudouridine synthase C

Similar to phosphoglycolate phosphatase, clustered with ubiquinone biosynthesis SAM-dependent O-methyltransferase

Similar to secretin RcpA/CpaC, associated with Flp pilus assembly

Similar to TadZ/CpaE, associated with Flp pilus assembly

Similarity with glutathionylspermidine synthase (EC 6.3.1.8), group 1

Single-stranded DNA-binding protein

Single-stranded-DNA-specific exonuclease RecJ (EC 3.1.-.-)

Sirohydrochlorin cobaltochelatase (EC 4.99.1.3)

Site-specific tyrosine recombinase

Sodium-Choline Symporter

Sodium-dependent phosphate transporter

soluble [2Fe-2S] ferredoxin

Soluble cytochrome b562

Soluble lytic murein transglycosylase precursor (EC 3.2.1.-)

Soluble pyridine nucleotide transhydrogenase (EC 1.6.1.1)

Sorbitol operon transcription regulator

Sorbitol-6-phosphate 2-dehydrogenase (EC 1.1.1.140)

SOS-response repressor and protease LexA (EC 3.4.21.88)

Spermidine N1-acetyltransferase (EC 2.3.1.57)

Spermidine Putrescine ABC transporter permease component PotB (TC 3.A.1.11.1)

Spermidine Putrescine ABC transporter permease component potC (TC_3.A.1.11.1)

Spermidine synthase (EC 2.5.1.16)

Squalene--hopene cyclase (EC 5.4.99.17)

Urease accessory protein UreF
Urease accessory protein UreG
Urease alpha subunit (EC 3.5.1.5)
Urease beta subunit (EC 3.5.1.5)
Urocanate hydratase (EC 4.2.1.49)
Uronate isomerase (EC 5.3.1.12)
Uroporphyrinogen-III methyltransferase (EC 2.1.1.107)
Valine--pyruvate aminotransferase (EC 2.6.1.66)
Various polyols ABC transporter, ATP-binding component
Various polyols ABC transporter, periplasmic substrate-binding protein
Various polyols ABC transporter, permease component 1
Various polyols ABC transporter, permease component 2
Vitamin B12 ABC transporter, B12-binding component BtuF
Vitamin B12 ABC transporter, permease component BtuC
Xanthine and CO dehydrogenases maturation factor, XdhC/CoxF family
Xanthine dehydrogenase, iron-sulfur cluster and FAD-binding subunit A (1.17.1.4)
Xanthine dehydrogenase, molybdenum binding subunit (EC 1.17.1.4)
Xanthine permease
Xanthine transporter,putative
Xanthine/uracil transporter
Xanthine/uracil/thiamine/ascorbate permease family protein
Xylonate dehydratase (EC 4.2.1.82)
Xylose ABC transporter, permease protein XylH
Xylulose kinase (EC 2.7.1.17)
Xylulose-5-phosphate phosphoketolase (EC 4.1.2.9)
YciL protein
YjbH-like, GTP pyrophosphokinase domain
YpfJ protein, zinc metalloprotease superfamily
Zinc ABC transporter, inner membrane permease protein ZnuB
Zinc ABC transporter, periplasmic-binding protein ZnuA

SSU ribosomal protein S10p (S20e)
SSU ribosomal protein S11p (S14e)
SSU ribosomal protein S12p (S23e)
SSU ribosomal protein S13p (S18e)
SSU ribosomal protein S15p (S13e)
SSU ribosomal protein S16p
SSU ribosomal protein S17p (S11e)
SSU ribosomal protein S18p
SSU ribosomal protein S19p (S15e)
SSU ribosomal protein S1p
SSU ribosomal protein S20p
SSU ribosomal protein S21p
SSU ribosomal protein S2p (SAe)
SSU ribosomal protein S3p (S3e)
SSU ribosomal protein S4p (S9e)
SSU ribosomal protein S5p (S2e)
SSU ribosomal protein S6p
SSU ribosomal protein S7p (S5e)
SSU ribosomal protein S8p (S15Ae)
SSU ribosomal protein S9p (S16e)
Stage V sporulation protein involved in spore cortex synthesis (SpoVR)
Starvation sensing protein RspA
Starvation sensing protein RspB
Substrate-specific component BioY of biotin ECF transporter
Substrate-specific component CbiM of cobalt ECF transporter
Substrate-specific component CbrT of predicted cobalamin ECF transporter
Substrate-specific component FolT of folate ECF transporter
Substrate-specific component YkoE of thiamin-regulated ECF transporter for HydroxyMethylPyrimidine
Succinate dehydrogenase cytochrome b subunit
Succinate dehydrogenase cytochrome b-556 subunit

Zinc uptake regulation protein ZUR

Zinc-regulated outer membrane receptor

Succinate dehydrogenase flavoprotein subunit (EC 1.3.99.1)

Succinate dehydrogenase hydrophobic membrane anchor protein

Succinate dehydrogenase iron-sulfur protein (EC 1.3.99.1)

Succinate-semialdehyde dehydrogenase [NADP+] (EC 1.2.1.79)

Succinyl-CoA ligase [ADP-forming] alpha chain (EC 6.2.1.5)

Succinyl-CoA ligase [ADP-forming] beta chain (EC 6.2.1.5)

Succinyl-CoA:3-ketoacid-coenzyme A transferase subunit A (EC 2.8.3.5)

Succinyl-CoA:3-ketoacid-coenzyme A transferase subunit B (EC 2.8.3.5)
Succinylarginine dihydrolase (EC 3.5.3.23)
Succinylglutamate desuccinylase (EC 3.5.1.96)

Succinylglutamic semialdehyde dehydrogenase (EC 1.2.1.71)
Succinylornithine transaminase (EC 2.6.1.81)
Sucrose phosphorylase (EC 2.4.1.7)
Sugar diacid utilization regulator SdaR

Sugar/maltose fermentation stimulation protein homolog
Sulfate adenylyltransferase subunit 1 (EC 2.7.7.4)
Sulfate adenylyltransferase subunit 2 (EC 2.7.7.4)
Sulfate and thiosulfate binding protein CysP

Sulfate and thiosulfate import ATP-binding protein CysA (EC 3.6.3.25)
Sulfate permease
Sulfate transport system permease protein CysT
Sulfate transport system permease protein CysW
Sulfate-binding protein Sbp

Sulfide dehydrogenase [flavocytochrome C] flavoprotein chain precursor (EC 1.8.2.-)

Sulfite reductase [NADPH] flavoprotein alpha-component (EC 1.8.1.2)

Sulfite reductase [NADPH] hemoprotein beta-component (EC 1.8.1.2)
sulfonate monooxygenase
Sulfur carrier protein adenylyltransferase ThiF
Sulfur carrier protein ThiS
Sulfur oxidation molybdopterin C protein
sulfur oxidation protein SoxA
Sulfur oxidation protein SoxB

Sulfur oxidation protein SoxY
Sulfur oxidation protein SoxZ

Superoxide dismutase [Cu-Zn] precursor (EC 1.15.1.1)
Superoxide dismutase [Fe] (EC 1.15.1.1)
Superoxide dismutase SodM-like protein ChrF

Survival protein SurA precursor (Peptidyl-prolyl cis-trans isomerase SurA) (EC 5.2.1.8)
Tagatose 1,6-bisphosphate aldolase (EC 4.1.2.40)
Tagatose-6-phosphate kinase (EC 2.7.1.144)
Tagatose-6-phosphate kinase AgaZ (EC 2.7.1.144)
Tagatose-6-phosphate kinase GatZ (EC 2.7.1.144)
Taurine transport ATP-binding protein TauB
Taurine transport system permease protein TauC
Taurine-binding periplasmic protein TauA

TcuA: flavoprotein used to oxidize tricarballylate to cis-aconitate

TcuB: works with TcuA to oxidize tricarballylate to cis-aconitate

TcuC: integral membrane protein used to transport tricarballylate across the cell membrane
Tetraacyldisaccharide 4'-kinase (EC 2.7.1.130)
Thiamin ABC transporter, ATPase component

Thiamin ABC transporter, substrate-binding component

Thiamin ABC transporter, transmembrane component

Thiamin-phosphate pyrophosphorylase (EC 2.5.1.3)

Thiaminase II (EC 3.5.99.2) involved in salvage of thiamin pyrimidine moiety
Thiamine biosynthesis protein thiI
Thiamine-monophosphate kinase (EC 2.7.4.16)
Thiazole biosynthesis protein ThiG
ThiJ/PfpI family protein
Thiol peroxidase, Tpx-type (EC 1.11.1.15)
Thiol:disulfide interchange protein DsbC
Thiol:disulfide oxidoreductase TlpA
Thioredoxin
Thioredoxin reductase (EC 1.8.1.9)

Thiosulfate sulfurtransferase, rhodanese (EC 2.8.1.1)

Threonine catabolic operon transcriptional activator TdcR

Threonine dehydratase (EC 4.3.1.19)

Threonine dehydratase biosynthetic (EC 4.3.1.19)

Threonine dehydratase, catabolic (EC 4.3.1.19)

Threonine dehydrogenase and related Zn-dependent dehydrogenases

Threonine synthase (EC 4.2.3.1)

Threonyl-tRNA synthetase (EC 6.1.1.3)

Thymidine kinase (EC 2.7.1.21)

Thymidine phosphorylase (EC 2.4.2.4)

Thymidylate kinase (EC 2.7.4.9)

Thymidylate synthase (EC 2.1.1.45)

tmRNA-binding protein SmpB

TniB NTP-binding protein

Tol biopolymer transport system, TolR protein

TolA protein

tolB protein precursor, periplasmic protein involved in the tonb-independent uptake of group A colicins

Toluene-4-monooxygenase, subunit TmoA

Toluene-4-monooxygenase, subunit TmoC

Toluene-4-monooxygenase, subunit TmoF

TonB-dependent hemin , ferrichrome receptor

TonB-dependent receptor

Topoisomerase IV subunit A (EC 5.99.1.-)

Topoisomerase IV subunit B (EC 5.99.1.-)

TPR domain protein, putative component of TonB system

TPR repeat containing exported protein

Transaldolase (EC 2.2.1.2)

Transamidase GatB domain protein

Transcription antitermination protein NusG

Transcription elongation factor GreA

Transcription elongation factor GreB

Transcription repressor of multidrug efflux pump acrAB operon, TetR (AcrR) family

Transcription termination factor Rho

Transcription termination protein NusA

Transcription termination protein NusB

Transcription-repair coupling factor

Transcriptional (co)regulator CytR

Transcriptional activator GadE
Transcriptional activator MetR
Transcriptional activator of acetoin dehydrogenase operon AcoR
Transcriptional activator of cad operon
Transcriptional activator of maltose regulon, MalT
Transcriptional activator protein LysR
Transcriptional activator RfaH
Transcriptional activatory protein CaiF
Transcriptional regulator BkdR of isoleucine and valine catabolism operon
Transcriptional regulator GabR of GABA utilization (GntR family with aminotransferase-like domain)
Transcriptional regulator HxlR, formaldehyde assimilation
Transcriptional regulator of catabolic arginine decarboxylase (adiA)
Transcriptional regulator of D-allose utilization, RpiR family
Transcriptional regulator of mannitol utilization, DeoR family protein
Transcriptional regulator of succinyl CoA synthetase operon
Transcriptional regulator PhnF
Transcriptional regulator YcjW, LacI family, possibly involved in maltodextrin utilization pathway
transcriptional regulator, Crp/Fnr family
Transcriptional regulator, IclR family
Transcriptional regulator, MerR family
Transcriptional regulatory protein basR/pmrA
Transcriptional regulatory protein CitB, DpiA
Transcriptional regulatory protein RtcR
Transcriptional repressor of the fructose operon, DeoR family
Transcriptional repressor protein TyrR
Transcriptional repressor UidR
Transketolase (EC 2.2.1.1)
Transketolase, C-terminal section (EC 2.2.1.1)
Translation elongation factor G
Translation elongation factor G-related protein
Translation elongation factor LepA
Translation elongation factor P

Translation elongation factor P-related protein
Translation elongation factor Ts
Translation elongation factor Tu
Translation initiation factor 1
Translation initiation factor 2
Translation initiation factor 3
Transmembrane component BioN of energizing module of biotin ECF transporter
Transmembrane component CbiQ of energizing module of cobalt ECF transporter
Transmembrane component CbrV of energizing module of predicted cobalamin ECF transporter
Transmembrane component YkoC of energizing module of thiamin-regulated ECF transporter for HydroxyMethylPyrimidine
Transport ATP-binding protein CydC
Transport ATP-binding protein CydCD
Transport ATP-binding protein CydD
Transposase YfcI
TRAP transporter solute receptor, unknown substrate 1
TRAP transporter, 4TM/12TM fusion protein, unknown substrate 1
TRAP-type transport system, periplasmic component, predicted N-acetylneuraminate-binding protein
TRAP-type transport system, small permease component, predicted N-acetylneuraminate transporter
Trehalase (EC 3.2.1.28)
Trehalase (EC 3.2.1.28)
Trehalose operon transcriptional repressor
Trehalose phosphorylase (EC 2.4.1.64)
Trehalose synthase (EC 5.4.99.16)
Trehalose-6-phosphate hydrolase (EC 3.2.1.93)
Trehalose-6-phosphate phosphatase (EC 3.1.3.12)
Tricarboxylate transport membrane protein TctA
Trimethylamine-N-oxide reductase (EC 1.6.6.9)
Triosephosphate isomerase (EC 5.3.1.1)
Triphosphoribosyl-dephospho-CoA synthetase (EC 2.7.8.25)
Trk system potassium uptake protein TrkA
Trk system potassium uptake protein TrkG

TrkA-N:Sodium/hydrogen exchanger

tRNA (5-methylaminomethyl-2-thiouridylate)-methyltransferase (EC 2.1.1.61)

tRNA (Guanine37-N1) -methyltransferase (EC 2.1.1.31)

tRNA (uracil(54)-C5)-methyltransferase (EC 2.1.1.35)

tRNA nucleotidyltransferase (EC 2.7.7.21) (EC 2.7.7.25)
tRNA pseudouridine 13 synthase (EC 4.2.1.-)
tRNA pseudouridine synthase A (EC 4.2.1.70)
tRNA pseudouridine synthase B (EC 4.2.1.70)
tRNA-dependent lipid II--amino acid ligase
tRNA-guanine transglycosylase (EC 2.4.2.29)
tRNA-i(6)A37 methylthiotransferase
tRNA-specific 2-thiouridylase MnmA
tRNA-specific adenosine-34 deaminase (EC 3.5.4.-)
tRNA:Cm32/Um32 methyltransferase
tRNA(Cytosine32)-2-thiocytidine synthetase
tRNA(Ile)-lysidine synthetase (EC 6.3.4.19)
Tryptophan 2-monooxygenase (EC 1.13.12.3)
Tryptophan 2,3-dioxygenase (EC 1.13.11.11)
Tryptophan synthase alpha chain (EC 4.2.1.20)
Tryptophan synthase beta chain (EC 4.2.1.20)
Tryptophanase (EC 4.1.99.1)
Tryptophanyl-tRNA synthetase (EC 6.1.1.2)

TsaB protein, required for threonylcarbamoyladenosine (t(6)A) formation in tRNA
TsaC protein (YrdC domain) required for threonylcarbamoyladenosine t(6)A37 modification in

tRNA

TsaD/Kae1/Qri7 protein, required for threonylcarbamoyladenosine t(6)A37 formation in tRNA

TsaE protein, required for threonylcarbamoyladenosine t(6)A37 formation in tRNA

tungsten-containing formate dehydrogenase alpha subunit

tungsten-containing formate dehydrogenase beta subunit
Twin-arginine translocation protein TatA
Twin-arginine translocation protein TatB
Twin-arginine translocation protein TatC

twitching motility protein PilJ
Twitching motility protein PilT
Two-component nitrogen fixation transcriptional regulator FixJ
Two-component oxygen-sensor histidine kinase FixL
Two-component response regulator CreC
Two-component sensor histidine kinase PleC
Two-component sensor PilS
Two-component system response regulator OmpR
Two-component system response regulator QseB
type 1 fimbrae adaptor subunit FimF
type 1 fimbriae anchoring protein FimD
Type cbb3 cytochrome oxidase biogenesis protein CcoG, involved in Cu oxidation
Type cbb3 cytochrome oxidase biogenesis protein CcoG, involved in Cu oxidation
Type I restriction-modification system, DNA-methyltransferase subunit M (EC 2.1.1.72)
Type I restriction-modification system, restriction subunit R (EC 3.1.21.3)
Type I restriction-modification system, specificity subunit S (EC 3.1.21.3)
Type I secretion outer membrane protein, TolC precursor
Type II secretory pathway, ATPase PulE/Tfp pilus assembly pathway, ATPase PilB
Type II/IV secretion system ATP hydrolase TadA/VirB11/CpaF, TadA subfamily
Type II/IV secretion system ATPase TadZ/CpaE, associated with Flp pilus assembly
Type II/IV secretion system protein TadC, associated with Flp pilus assembly
Type II/IV secretion system secretin RcpA/CpaC, associated with Flp pilus assembly
Type III restriction-modification system methylation subunit (EC 2.1.1.72)
Type IV fimbriae expression regulatory protein PilR
Type IV fimbrial assembly protein PilC
Type IV fimbrial assembly, ATPase PilB
Type IV fimbrial biogenesis protein PilY1
type IV pili signal transduction protein PilI
Type IV pilin PilA
Type IV pilus biogenesis protein PilE
Type IV pilus biogenesis protein PilQ
Type IV prepilin peptidase TadV/CpaA

Type IV secretion system protein VirD4

Tyrosine recombinase XerC

Tyrosine-protein kinase Wzc (EC 2.7.10.2)

Tyrosyl-tRNA synthetase (EC 6.1.1.1)

ubiquinol cytochrome C oxidoreductase, cytochrome C1 subunit

Ubiquinol--cytochrome c reductase, cytochrome B subunit (EC 1.10.2.2)

Ubiquinol-cytochrome C reductase iron-sulfur subunit (EC 1.10.2.2)

Ubiquinone biosynthesis monooxygenase UbiB

Ubiquinone/menaquinone biosynthesis methyltransferase UbiE (EC 2.1.1.-)

UDP-3-O-[3-hydroxymyristoyl] N-acetylglucosamine deacetylase (EC 3.5.1.108)

UDP-4-amino-4-deoxy-L-arabinose formyltransferase (EC 2.1.2.13)

UDP-4-amino-4-deoxy-L-arabinose--oxoglutarate aminotransferase (EC 2.6.1.-)

UDP-galactopyranose mutase (EC 5.4.99.9)

UDP-glucose 4-epimerase (EC 5.1.3.2)

UDP-glucose dehydrogenase (EC 1.1.1.22)

UDP-glucose:(heptosyl) LPS alpha1,3-glucosyltransferase WaaG (EC 2.4.1.-)

UDP-glucuronic acid oxidase (UDP-4-keto-hexauronic acid decarboxylating) (EC 1.1.1.305)

UDP-N-acetylenolpyruvoylglucosamine reductase (EC 1.1.1.158)

UDP-N-acetylglucosamine 1-carboxyvinyltransferase (EC 2.5.1.7)

UDP-N-acetylglucosamine 2-epimerase (EC 5.1.3.14)

UDP-N-acetylglucosamine 4,6-dehydratase (EC 4.2.1.-)

UDP-N-acetylglucosamine--N-acetylmuramyl-(pentapeptide) pyrophosphoryl-undecaprenol N-acetylglucosamine transferase (EC 2.4.1.227)

UDP-N-acetylmuramate--alanine ligase (EC 6.3.2.8)

UDP-N-acetylmuramoylalanine--D-glutamate ligase (EC 6.3.2.9)

UDP-N-acetylmuramoylalanyl-D-glutamate--2,6-diaminopimelate ligase (EC 6.3.2.13)

UDP-N-acetylmuramoylalanyl-D-glutamyl-2,6-diaminopimelate--D-alanyl-D-alanine ligase (EC 6.3.2.10)

Uncharacterized ABC transporter, periplasmic component YrbD

Uncharacterized glutathione S-transferase-like protein

Uncharacterized GST-like protein yghU associated with glutathionylspermidine synthetase/amidase

Uncharacterized GST-like protein yibF
Uncharacterized MobA-related protein
Uncharacterized monothiol glutaredoxin ycf64-like
Uncharacterized protein ImpA
Uncharacterized protein ImpJ/VasE
Uncharacterized protein YadU in stf fimbrial cluster
Uncharacterized protein YidS
Uncharacterized protein YphG, TPR-domain containing
Uncharacterized protein, similar to the N-terminal domain of Lon protease
Uncharacterized sugar kinase YdjH
Uncharacterized sugar kinase YeiI
Undecaprenyl-phosphate galactosephosphotransferase (EC 2.7.8.6)
Undecaprenyl-phosphate N-acetylglucosaminyl 1-phosphate transferase (EC 2.7.8.-)
Universal stress protein E
Universal stress protein F
Universal stress protein family
Unsaturated fatty acid biosythesis repressor FabR, TetR family
UPF0141 membrane protein YijP possibly required for phosphoethanolamine modification of lipopolysaccharide
UPF0234 protein YajQ
Uptake hydrogenase large subunit (EC 1.12.99.6)
Uptake hydrogenase small subunit precursor (EC 1.12.99.6)
Uracil permease
Uracil phosphoribosyltransferase (EC 2.4.2.9)
Uracil-DNA glycosylase, family 1
Uracil-DNA glycosylase, family 4
Uracil-DNA glycosylase, putative family 6
Urea ABC transporter, ATPase protein UrtD
Urea ABC transporter, ATPase protein UrtE
Urea ABC transporter, permease protein UrtB
Urea ABC transporter, permease protein UrtC
Urea ABC transporter, urea binding protein
Urea carboxylase (EC 6.3.4.6)
Urea carboxylase-related ABC transporter, ATPase protein

Urea carboxylase-related ABC transporter, periplasmic substrate-binding protein
Urea carboxylase-related ABC transporter, permease protein
Urea carboxylase-related aminomethyltransferase (EC 2.1.2.10)
Urease accessory protein UreD
Urease accessory protein UreE
Urease accessory protein UreG
Urease alpha subunit (EC 3.5.1.5)
Ureidoglycolate dehydrogenase (EC 1.1.1.154)
Ureidoglycolate hydrolase (EC 3.5.3.19)
Urocanate hydratase (EC 4.2.1.49)
Uronate isomerase (EC 5.3.1.12)
Uroporphyrinogen-III methyltransferase (EC 2.1.1.107)
Uxu operon transcriptional regulator
V-type ATP synthase subunit B (EC 3.6.3.14)
V-type ATP synthase subunit E (EC 3.6.3.14)
Valine--pyruvate aminotransferase (EC 2.6.1.66)
VapC toxin protein
Various polyols ABC transporter, ATP-binding component
Various polyols ABC transporter, periplasmic substrate-binding protein
Various polyols ABC transporter, permease component 1
Very-short-patch mismatch repair endonuclease (G-T specific)
Vitamin B12 ABC transporter, ATPase component BtuD
Vitamin B12 ABC transporter, B12-binding component BtuF
Vitamin B12 ABC transporter, permease component BtuC
Xaa-Pro dipeptidase PepQ (EC 3.4.13.9)
Xanthine and CO dehydrogenases maturation factor, XdhC/CoxF family
Xanthine dehydrogenase, FAD binding subunit (EC 1.17.1.4)
Xanthine dehydrogenase, iron-sulfur cluster and FAD-binding subunit A (1.17.1.4)
Xanthine dehydrogenase, molybdenum binding subunit (EC 1.17.1.4)
Xanthine permease
Xanthine transporter,putative
Xanthine-guanine phosphoribosyltransferase (EC 2.4.2.22)

Xanthine/uracil transporter
Xanthine/uracil/thiamine/ascorbate permease family protein
Xanthosine permease
Xanthosine phosphorylase (EC 2.4.2.1)
XdhC protein (assists in molybdopterin insertion into xanthine dehydrogenase)
Xylonate dehydratase (EC 4.2.1.82)
Xylose ABC transporter, periplasmic xylose-binding protein XylF
Xylose activator XylR (AraC family)
Xyloside transporter XynT
Xylulose kinase (EC 2.7.1.17)
Xylulose-5-phosphate phosphoketolase (EC 4.1.2.9)
YbbL ABC transporter ATP-binding protein
YbbM seven transmembrane helix protein
YefM protein (antitoxin to YoeB)
YjiD protein
YpfJ protein, zinc metalloprotease superfamily
Zinc ABC transporter, ATP-binding protein ZnuC
Zinc ABC transporter, inner membrane permease protein ZnuB
Zinc ABC transporter, periplasmic-binding protein ZnuA
Zinc metalloproteinase precursor (EC 3.4.24.29)
Zinc resistance-associated protein
Zinc uptake regulation protein ZUR
Zinc-type alcohol dehydrogenase YcjQ

Appendix 4: List of total predicted genes in HsAqS

LADP5 (LTW0004)

(R)-citramalate synthase (EC 2.3.1.182)
1-hydroxy-2-methyl-2-(E)-butenyl 4-diphosphate synthase (EC 1.17.7.1)
1,4-alpha-glucan (glycogen) branching enzyme, GH-13-type (EC 2.4.1.18)
2-amino-3-ketobutyrate coenzyme A ligase (EC 2.3.1.29)
2-C-methyl-D-erythritol 2,4-cyclodiphosphate synthase (EC 4.6.1.12)
2-dehydro-3-deoxyphosphogluconate aldolase (EC 4.1.2.14)
2-hydroxy-6-oxo-6-phenylhexa-2,4-dienoate hydrolase (EC 3.7.1.-)
2-hydroxyhepta-2,4-diene-1,7-dioate isomerase (EC 5.3.3.-)
2-isopropylmalate synthase (EC 2.3.3.13)
2-keto-3-deoxy-D-arabino-heptulosonate-7-phosphate synthase I beta (EC 2.5.1.54)
2-keto-3-deoxy-D-arabino-heptulosonate-7-phosphate synthase II (EC 2.5.1.54)
2-methylcitrate dehydratase (EC 4.2.1.79)
2-succinyl-5-enolpyruvyl-6-hydroxy-3-cyclohexene-1-carboxylic-acid synthase (EC 2.2.1.9)
2,3-bisphosphoglycerate-independent phosphoglycerate mutase (EC 5.4.2.1)
2,3,4,5-tetrahydropyridine-2,6-dicarboxylate N-succinyltransferase (EC 2.3.1.117)
3-dehydroquinate dehydratase II (EC 4.2.1.10)
3-demethylubiquinone-9 3-methyltransferase (EC 2.1.1.64)
3-hydroxyacyl-CoA dehydrogenase (EC 1.1.1.35)

LADP6 (LTW0005)

(R)-citramalate synthase (EC 2.3.1.182)
(S)-2,3-di-O-geranylgeranylglyceryl phosphate synthase
(S)-3-O-geranylgeranylglyceryl phosphate synthase
1-deoxy-D-xylulose 5-phosphate reductoisomerase (EC 1.1.1.267)
1-deoxy-D-xylulose 5-phosphate synthase (EC 2.2.1.7)
1-hydroxy-2-methyl-2-(E)-butenyl 4-diphosphate synthase (EC 1.17.7.1)
1,2-dihydroxy-3-keto-5-methylthiopentene dioxygenase (EC 1.13.11.54)
1,4-alpha-glucan (glycogen) branching enzyme, GH-13-type (EC 2.4.1.18)
2-amino-3-ketobutyrate coenzyme A ligase (EC 2.3.1.29)
2-amino-3,7-dideoxy-D-threo-hept-6-ulosonate synthase (EC 2.2.1.10)
2-dehydro-3-deoxyphosphogluconate aldolase (EC 4.1.2.14)
2-dehydropantoate 2-reductase (EC 1.1.1.169)
2-hydroxy-3-oxopropionate reductase (EC 1.1.1.60)
2-hydroxy-6-oxo-6-phenylhexa-2,4-dienoate hydrolase (EC 3.7.1.-)
2-hydroxyhepta-2,4-diene-1,7-dioate isomerase (EC 5.3.3.-)
2-isopropylmalate synthase (EC 2.3.3.13)
2-keto-3-deoxy-D-arabino-heptulosonate-7-phosphate synthase I beta (EC 2.5.1.54)
2-keto-3-deoxy-D-arabino-heptulosonate-7-phosphate

LA (LTW0006)

(R)-citramalate synthase (EC 2.3.1.182)
(S)-2,3-di-O-geranylgeranylglyceryl phosphate synthase
(S)-3-O-geranylgeranylglyceryl phosphate synthase
1-deoxy-D-xylulose 5-phosphate reductoisomerase (EC 1.1.1.267)
1-deoxy-D-xylulose 5-phosphate synthase (EC 2.2.1.7)
1-hydroxy-2-methyl-2-(E)-butenyl 4-diphosphate synthase (EC 1.17.7.1)
1,2-dihydroxy-3-keto-5-methylthiopentene dioxygenase (EC 1.13.11.54)
1,4-alpha-glucan (glycogen) branching enzyme, GH-13-type (EC 2.4.1.18)
2-amino-3-ketobutyrate coenzyme A ligase (EC 2.3.1.29)
2-amino-3,7-dideoxy-D-threo-hept-6-ulosonate synthase (EC 2.2.1.10)
2-dehydro-3-deoxyphosphogluconate aldolase (EC 4.1.2.14)
2-dehydropantoate 2-reductase (EC 1.1.1.169)
2-hydroxy-3-oxopropionate reductase (EC 1.1.1.60)
2-hydroxy-6-oxo-6-phenylhexa-2,4-dienoate hydrolase (EC 3.7.1.-)
2-hydroxyhepta-2,4-diene-1,7-dioate isomerase (EC 5.3.3.-)
2-isopropylmalate synthase (EC 2.3.3.13)
2-keto-3-deoxy-D-arabino-heptulosonate-7-phosphate synthase I beta (EC 2.5.1.54)
2-keto-3-deoxy-D-arabino-heptulosonate-7-phosphate synthase II

MUP (LTW0007)

(R)-citramalate synthase (EC 2.3.1.182)
(S)-3-O-geranylgeranylglyceryl phosphate synthase
[NiFe] hydrogenase metallocenter assembly protein HypD
[NiFe] hydrogenase metallocenter assembly protein HypF
1,4-alpha-glucan (glycogen) branching enzyme, GH-13-type (EC 2.4.1.18)
2-amino-3-ketobutyrate coenzyme A ligase (EC 2.3.1.29)
2-C-methyl-D-erythritol 2,4-cyclodiphosphate synthase (EC 4.6.1.12)
2-hydroxyhepta-2,4-diene-1,7-dioate isomerase (EC 5.3.3.-)
2-isopropylmalate synthase (EC 2.3.3.13)
2-keto-3-deoxy-D-arabino-heptulosonate-7-phosphate synthase I beta (EC 2.5.1.54)
2-keto-3-deoxy-D-arabino-heptulosonate-7-phosphate synthase II (EC 2.5.1.54)
2-methylcitrate dehydratase (EC 4.2.1.79)
2,3-bisphosphoglycerate-independent phosphoglycerate mutase (EC 5.4.2.1)
3-dehydroquinate dehydratase I (EC 4.2.1.10)
3-hydroxyacyl-CoA dehydrogenase (EC 1.1.1.35)
3-hydroxybutyryl-CoA dehydrogenase (EC 1.1.1.157)
3-isopropylmalate dehydratase large subunit (EC 4.2.1.33)
3-isopropylmalate dehydratase small subunit (EC 4.2.1.33)

3-hydroxybutyryl-CoA dehydrogenase (EC 1.1.1.157)
3-isopropylmalate dehydratase large subunit (EC 4.2.1.33)
3-isopropylmalate dehydratase small subunit (EC 4.2.1.33)
3-ketoacyl-CoA thiolase (EC 2.3.1.16)
3-methyl-2-oxobutanoate hydroxymethyltransferase (EC 2.1.2.11)
3-oxoacyl-[acyl-carrier protein] reductase (EC 1.1.1.100)
3,4-dihydroxy-2-butanone 4-phosphate synthase (EC 4.1.99.12)
3,4-dihydroxyphenylacetate 2,3-dioxygenase (EC 1.13.11.15)
3'-to-5' exoribonuclease RNase R
4-alpha-glucanotransferase (amylomaltase) (EC 2.4.1.25)
4-diphosphocytidyl-2-C-methyl-D-erythritol kinase (EC 2.7.1.148)
4-hydroxy-3-methylbut-2-enyl diphosphate reductase (EC 1.17.1.2)
4-hydroxyphenylacetate 3-monooxygenase (EC 1.14.13.3)
4-hydroxyproline epimerase (EC 5.1.1.8)
4-hydroxythreonine-4-phosphate dehydrogenase (EC 1.1.1.262)
5-carboxymethyl uridine and 5-carboxymethyl 2-thiouridine methyltransferase
5-carboxymethyl-2-hydroxymuconate semialdehyde dehydrogenase (EC 1.2.1.60)
5-Enolpyruvylshikimate-3-phosphate synthase (EC 2.5.1.19)
5-FCL-like protein

synthase II (EC 2.5.1.54)
2-succinyl-5-enolpyruvyl-6-hydroxy-3-cyclohexene-1-carboxylic-acid synthase (EC 2.2.1.9)
2-succinyl-6-hydroxy-2,4-cyclohexadiene-1-carboxylate synthase (EC 4.2.99.20)
2,3-bisphosphoglycerate-independent phosphoglycerate mutase (EC 5.4.2.1)
2,3,4,5-tetrahydropyridine-2,6-dicarboxylate N-acetyltransferase (EC 2.3.1.89)
2,3,4,5-tetrahydropyridine-2,6-dicarboxylate N-succinyltransferase (EC 2.3.1.117)
3-dehydroquinate dehydratase II (EC 4.2.1.10)
3-dehydroquinate synthase (EC 4.2.3.4)
3-demethylubiquinone-9 3-methyltransferase (EC 2.1.1.64)
3-hydroxyacyl-[acyl-carrier-protein] dehydratase, FabZ form (EC 4.2.1.59)
3-hydroxyacyl-CoA dehydrogenase (EC 1.1.1.35)
3-hydroxybutyryl-CoA dehydrogenase (EC 1.1.1.157)
3-isopropylmalate dehydratase large subunit (EC 4.2.1.33)
3-isopropylmalate dehydrogenase (EC 1.1.1.85)
3-ketoacyl-CoA thiolase (EC 2.3.1.16)
3-oxoacyl-[acyl-carrier protein] reductase (EC 1.1.1.100)
3,4-dihydroxy-2-butanone 4-phosphate synthase (EC 4.1.99.12)
3'-to-5' exoribonuclease RNase R
4-alpha-glucanotransferase (amylomaltase) (EC 2.4.1.25)
4-diphosphocytidyl-2-C-methyl-D-erythritol kinase (EC

(EC 2.5.1.54)
2-succinyl-5-enolpyruvyl-6-hydroxy-3-cyclohexene-1-carboxylic-acid synthase (EC 2.2.1.9)
2-succinyl-6-hydroxy-2,4-cyclohexadiene-1-carboxylate synthase (EC 4.2.99.20)
2,3-bisphosphoglycerate-independent phosphoglycerate mutase (EC 5.4.2.1)
2,3,4,5-tetrahydropyridine-2,6-dicarboxylate N-acetyltransferase (EC 2.3.1.89)
2,3,4,5-tetrahydropyridine-2,6-dicarboxylate N-succinyltransferase (EC 2.3.1.117)
3-dehydroquinate dehydratase II (EC 4.2.1.10)
3-dehydroquinate synthase (EC 4.2.3.4)
3-demethylubiquinone-9 3-methyltransferase (EC 2.1.1.64)
3-hydroxyacyl-[acyl-carrier-protein] dehydratase, FabZ form (EC 4.2.1.59)
3-hydroxyacyl-CoA dehydrogenase (EC 1.1.1.35)
3-hydroxybutyryl-CoA dehydrogenase (EC 1.1.1.157)
3-isopropylmalate dehydratase large subunit (EC 4.2.1.33)
3-isopropylmalate dehydrogenase (EC 1.1.1.85)
3-ketoacyl-CoA thiolase (EC 2.3.1.16)
3-oxoacyl-[acyl-carrier protein] reductase (EC 1.1.1.100)
3,4-dihydroxy-2-butanone 4-phosphate synthase (EC 4.1.99.12)
3'-to-5' exoribonuclease RNase R
4-alpha-glucanotransferase (amylomaltase) (EC 2.4.1.25)
4-diphosphocytidyl-2-C-methyl-D-erythritol kinase (EC 2.7.1.148)

3-isopropylmalate dehydrogenase (EC 1.1.1.85)
3-ketoacyl-CoA thiolase (EC 2.3.1.16)
3-methyl-2-oxobutanoate hydroxymethyltransferase (EC 2.1.2.11)
3-oxoacyl-[acyl-carrier protein] reductase (EC 1.1.1.100)
3,4-dihydroxyphenylacetate 2,3-dioxygenase (EC 1.13.11.15)
3,7-dideoxy-D-threo-hepto-2, 6-diulosonate synthase (EC 1.4.1.24)
4-hydroxy-3-methylbut-2-enyl diphosphate reductase (EC 1.17.1.2)
4-hydroxyphenylacetate 3-monooxygenase (EC 1.14.13.3)
4-hydroxythreonine-4-phosphate dehydrogenase (EC 1.1.1.262)
5-Enolpyruvylshikimate-3-phosphate synthase (EC 2.5.1.19)
5-FCL-like protein
5-methyltetrahydropteroyltriglutamate--homocysteine methyltransferase (EC 2.1.1.14)
5-nucleotidase SurE (EC 3.1.3.5)
5,10-methylenetetrahydrofolate reductase (EC 1.5.1.20)
6-phosphofructokinase (EC 2.7.1.11)
6-phosphogluconate dehydrogenase, decarboxylating (EC 1.1.1.44)
6-phosphogluconolactonase (EC 3.1.1.31), eukaryotic type
6,7-dimethyl-8-ribityllumazine synthase (EC 2.5.1.78)
7,8-didemethyl-8-hydroxy-5-deazariboflavin synthase subunit 1

5-FCL-like protein, but predicted not to be 5-formyltetrahydrofolate cyclo-ligase (5-FCL)
5-methyltetrahydropteroyltriglutamate--homocysteine methyltransferase (EC 2.1.1.14)
5,10-methylenetetrahydrofolate reductase (EC 1.5.1.20)
5'-methylthioadenosine phosphorylase (EC 2.4.2.28)
5'-nucleotidase (EC 3.1.3.5)
6-phosphofructokinase (EC 2.7.1.11)
6-phosphogluconate dehydrogenase, decarboxylating (EC 1.1.1.44)
6-pyruvoyl tetrahydrobiopterin synthase (EC 4.2.3.12)
6,7-dimethyl-8-ribityllumazine synthase (EC 2.5.1.78)
7,8-didemethyl-8-hydroxy-5-deazariboflavin synthase subunit 1
7,8-didemethyl-8-hydroxy-5-deazariboflavin synthase subunit 2
A/G-specific adenine glycosylase (EC 3.2.2.-)
ABC transporter, periplasmic spermidine putrescine-binding protein PotD (TC 3.A.1.11.1)
ABC-type nitrate/sulfonate/bicarbonate transport system, permease component
Acetaldehyde dehydrogenase (EC 1.2.1.10)
Acetolactate synthase large subunit (EC 2.2.1.6)
Acetone carboxylase, alpha subunit (EC 6.4.1.6)
Acetone carboxylase, beta subunit (EC 6.4.1.6)
Acetone carboxylase, gamma subunit (EC 6.4.1.6)

2.7.1.148)
4-hydroxy-3-methylbut-2-enyl diphosphate reductase (EC 1.17.1.2)
4-hydroxyproline epimerase (EC 5.1.1.8)
4-hydroxythreonine-4-phosphate dehydrogenase (EC 1.1.1.262)
5-carboxymethyl-2-hydroxymuconate semialdehyde dehydrogenase (EC 1.2.1.60)
5-Enolpyruvylshikimate-3-phosphate synthase (EC 2.5.1.19)
5-FCL-like protein
5-FCL-like protein, but predicted not to be 5-formyltetrahydrofolate cyclo-ligase (5-FCL)
5-formyltetrahydrofolate cyclo-ligase (EC 6.3.3.2)
5-keto-D-gluconate 5-reductase (EC 1.1.1.69)
5-methyltetrahydropteroyltriglutamate--homocysteine methyltransferase (EC 2.1.1.14)
5-nucleotidase SurE (EC 3.1.3.5)
5,10-methylenetetrahydrofolate reductase (EC 1.5.1.20)
5'-methylthioadenosine nucleosidase (EC 3.2.2.16)
5'-methylthioadenosine phosphorylase (EC 2.4.2.28)
5'-nucleotidase (EC 3.1.3.5)
5'-nucleotidase family protein in cluster with NagD-like phosphatase
6-phosphofructokinase (EC 2.7.1.11)
6-phosphogluconate dehydrogenase, decarboxylating (EC 1.1.1.44)
A/G-specific adenine glycosylase (EC 3.2.2.-)

4-hydroxy-3-methylbut-2-enyl diphosphate reductase (EC 1.17.1.2)
4-hydroxyproline epimerase (EC 5.1.1.8)
4-hydroxythreonine-4-phosphate dehydrogenase (EC 1.1.1.262)
5-carboxymethyl-2-hydroxymuconate semialdehyde dehydrogenase (EC 1.2.1.60)
5-Enolpyruvylshikimate-3-phosphate synthase (EC 2.5.1.19)
5-FCL-like protein
5-FCL-like protein, but predicted not to be 5-formyltetrahydrofolate cyclo-ligase (5-FCL)
5-formyltetrahydrofolate cyclo-ligase (EC 6.3.3.2)
5-keto-D-gluconate 5-reductase (EC 1.1.1.69)
5-methyltetrahydropteroyltriglutamate--homocysteine methyltransferase (EC 2.1.1.14)
5-nucleotidase SurE (EC 3.1.3.5)
5,10-methylenetetrahydrofolate reductase (EC 1.5.1.20)
5'-methylthioadenosine nucleosidase (EC 3.2.2.16)
5'-methylthioadenosine phosphorylase (EC 2.4.2.28)
5'-nucleotidase (EC 3.1.3.5)
5'-nucleotidase family protein in cluster with NagD-like phosphatase
6-phosphofructokinase (EC 2.7.1.11)
6-phosphogluconate dehydrogenase, decarboxylating (EC 1.1.1.44)
A/G-specific adenine glycosylase (EC 3.2.2.-)

ABC-type nitrate/sulfonate/bicarbonate transport system, permease component
Acetoin dehydrogenase E1 component alpha-subunit (EC 1.2.4.-)
Acetoin dehydrogenase E1 component beta-subunit (EC 1.2.4.-)
Acetolactate synthase large subunit (EC 2.2.1.6)
Acetone carboxylase, alpha subunit (EC 6.4.1.6)
Acetone carboxylase, beta subunit (EC 6.4.1.6)
Acetone carboxylase, gamma subunit (EC 6.4.1.6)
Acetyl-CoA acetyltransferase (EC 2.3.1.9)
Acetyl-CoA synthetase (ADP-forming) alpha and beta chains, putative
Acetyl-coenzyme A carboxyl transferase alpha chain (EC 6.4.1.2)
Acetyl-coenzyme A carboxyl transferase beta chain (EC 6.4.1.2)
Acetyl-coenzyme A synthetase (EC 6.2.1.1)
Acetylornithine aminotransferase (EC 2.6.1.11)
Acetylornithine deacetylase (EC 3.5.1.16)
Acetylornithine deacetylase/Succinyl-diaminopimelate desuccinylase and related deacylases
Aconitate hydratase (EC 4.2.1.3)
Acriflavin resistance protein
Acyl carrier protein
Acylamino-acid-releasing enzyme (EC 3.4.19.1)

Acetyl-coenzyme A carboxyl transferase alpha chain (EC 6.4.1.2)
Acetyl-coenzyme A carboxyl transferase beta chain (EC 6.4.1.2)
Acetyl-coenzyme A synthetase (EC 6.2.1.1)
Acetylglutamate kinase (EC 2.7.2.8)
Acetylornithine aminotransferase (EC 2.6.1.11)
Aconitate hydratase (EC 4.2.1.3)
Acriflavin resistance protein
Acyl-phosphate:glycerol-3-phosphate O-acyltransferase PlsY
Adenine deaminase (EC 3.5.4.2)
Adenosine deaminase (EC 3.5.4.4)
Adenosylhomocysteinase (EC 3.3.1.1)
Adenylate kinase (EC 2.7.4.3)
Adenylosuccinate lyase (EC 4.3.2.2)
Adenylosuccinate synthetase (EC 6.3.4.4)
Adenylylsulfate kinase (EC 2.7.1.25)
ADP-ribose pyrophosphatase (EC 3.6.1.13)
Aerobic glycerol-3-phosphate dehydrogenase (EC 1.1.5.3)
Agmatinase (EC 3.5.3.11)
Alanine dehydrogenase (EC 1.4.1.1)
Alanyl-tRNA synthetase (EC 6.1.1.7)
Alcohol dehydrogenase (EC 1.1.1.1)

AA3-600 quinol oxidase subunit I
AA3-600 quinol oxidase subunit II
ABC transporter, periplasmic spermidine putrescine-binding protein PotD (TC 3.A.1.11.1)
ABC-type nitrate/sulfonate/bicarbonate transport system, permease component
Acetaldehyde dehydrogenase (EC 1.2.1.10)
Acetate kinase (EC 2.7.2.1)
Acetate permease ActP (cation/acetate symporter)
Acetoacetyl-CoA reductase (EC 1.1.1.36)
Acetoin dehydrogenase E1 component alpha-subunit (EC 1.2.4.-)
Acetoin dehydrogenase E1 component beta-subunit (EC 1.2.4.-)
Acetolactate synthase large subunit (EC 2.2.1.6)
Acetone carboxylase, alpha subunit (EC 6.4.1.6)
Acetone carboxylase, beta subunit (EC 6.4.1.6)
Acetyl-CoA acetyltransferase (EC 2.3.1.9)
Acetyl-CoA synthetase (ADP-forming) alpha and beta chains, putative
Acetyl-CoA:acetoacetyl-CoA transferase, beta subunit (EC 2.8.3.8)
Acetyl-coenzyme A carboxyl transferase alpha chain (EC 6.4.1.2)
Acetyl-coenzyme A carboxyl transferase beta chain (EC 6.4.1.2)
Acetyl-coenzyme A synthetase (EC 6.2.1.1)
Acetylglutamate kinase (EC 2.7.2.8)
Acetylornithine aminotransferase (EC 2.6.1.11)

AA3-600 quinol oxidase subunit I
AA3-600 quinol oxidase subunit II
ABC transporter, periplasmic spermidine putrescine-binding protein PotD (TC 3.A.1.11.1)
ABC-type nitrate/sulfonate/bicarbonate transport system, permease component
Acetaldehyde dehydrogenase (EC 1.2.1.10)
Acetate kinase (EC 2.7.2.1)
Acetate permease ActP (cation/acetate symporter)
Acetoacetyl-CoA reductase (EC 1.1.1.36)
Acetoin dehydrogenase E1 component alpha-subunit (EC 1.2.4.-)
Acetoin dehydrogenase E1 component beta-subunit (EC 1.2.4.-)
Acetolactate synthase large subunit (EC 2.2.1.6)
Acetone carboxylase, alpha subunit (EC 6.4.1.6)
Acetone carboxylase, beta subunit (EC 6.4.1.6)
Acetyl-CoA acetyltransferase (EC 2.3.1.9)
Acetyl-CoA synthetase (ADP-forming) alpha and beta chains, putative
Acetyl-CoA:acetoacetyl-CoA transferase, beta subunit (EC 2.8.3.8)
Acetyl-coenzyme A carboxyl transferase alpha chain (EC 6.4.1.2)
Acetyl-coenzyme A carboxyl transferase beta chain (EC 6.4.1.2)
Acetyl-coenzyme A synthetase (EC 6.2.1.1)
Acetylglutamate kinase (EC 2.7.2.8)
Acetylornithine aminotransferase (EC 2.6.1.11)

Adenine phosphoribosyltransferase (EC 2.4.2.7)
Adenosylhomocysteinase (EC 3.3.1.1)
Adenylate kinase (EC 2.7.4.3)
Adenylosuccinate lyase (EC 4.3.2.2)
Adenylosuccinate synthetase (EC 6.3.4.4)
Adenylylsulfate kinase (EC 2.7.1.25)
ADP-ribose pyrophosphatase (EC 3.6.1.13)
Aerobic glycerol-3-phosphate dehydrogenase (EC 1.1.5.3)
Alanine dehydrogenase (EC 1.4.1.1)
Alanyl-tRNA synthetase (EC 6.1.1.7)
Alcohol dehydrogenase (EC 1.1.1.1)
Aldehyde dehydrogenase (EC 1.2.1.3)
Aldehyde dehydrogenase B (EC 1.2.1.22)
Alkaline phosphatase (EC 3.1.3.1)
Alternative cytochrome c oxidase polypeptide CoxN (EC 1.9.3.1)
Alternative dihydrofolate reductase 2
Amidophosphoribosyltransferase (EC 2.4.2.14)
Aminodeoxychorismate lyase (EC 4.1.3.38)
Aminomethyltransferase (glycine cleavage system T protein) (EC 2.1.2.10)
Anaerobic glycerol-3-phosphate dehydrogenase subunit A (EC 1.1.5.3)
Anaerobic glycerol-3-phosphate dehydrogenase subunit C (EC 1.1.5.3)

Aldehyde dehydrogenase (EC 1.2.1.3)
Aldehyde dehydrogenase B (EC 1.2.1.22)
Aldose 1-epimerase (EC 5.1.3.3)
Alfa-L-rhamnosidase (EC 3.2.1.40)
Alkaline phosphatase (EC 3.1.3.1)
Alkyl hydroperoxide reductase subunit C-like protein
Alpha-1,2-mannosidase
Alpha-arabinosides ABC transport system, permease protein 1
Alpha-glucoside transport system permease protein AglG
Alpha,alpha-trehalose-phosphate synthase [UDP-forming] (EC 2.4.1.15)
Alternative cytochrome c oxidase polypeptide CoxN (EC 1.9.3.1)
Amidophosphoribosyltransferase (EC 2.4.2.14)
Aminodeoxychorismate lyase (EC 4.1.3.38)
Aminomethyltransferase (glycine cleavage system T protein) (EC 2.1.2.10)
Anaerobic glycerol-3-phosphate dehydrogenase subunit A (EC 1.1.5.3)
Anaerobic glycerol-3-phosphate dehydrogenase subunit C (EC 1.1.5.3)
Anthranilate phosphoribosyltransferase (EC 2.4.2.18)
Anthranilate synthase, amidotransferase component (EC 4.1.3.27)
Arabinan endo-1,5-alpha-L-arabinosidase (EC 3.2.1.99)
archaeal ATPase, fused to C-terminal DUF234 domain
Archaeal DNA polymerase I (EC 2.7.7.7)
Archaeal DNA polymerase II large subunit (EC 2.7.7.7)

Acetylornithine deacetylase (EC 3.5.1.16)
Acetyltransferase AcuA, acetyl-CoA synthetase inhibitor
Aconitate hydratase (EC 4.2.1.3)
Acyl carrier protein
Acyl-phosphate:glycerol-3-phosphate O-acyltransferase PlsY
Acylphosphate phosphohydrolase (EC 3.6.1.7), putative
Adenine deaminase (EC 3.5.4.2)
Adenine phosphoribosyltransferase (EC 2.4.2.7)
Adenosine deaminase (EC 3.5.4.4)
Adenosylcobinamide-phosphate guanylyltransferase (EC 2.7.7.62)
Adenosylhomocysteinase (EC 3.3.1.1)
Adenosylmethionine-8-amino-7-oxononanoate aminotransferase (EC 2.6.1.62)
Adenylate cyclase (EC 4.6.1.1)
Adenylate kinase (EC 2.7.4.3)
Adenylosuccinate lyase (EC 4.3.2.2)
Adenylosuccinate synthetase (EC 6.3.4.4)
Adenylylsulfate kinase (EC 2.7.1.25)
ADP-ribose pyrophosphatase (EC 3.6.1.13)
Aerobic cobaltochelatase CobS subunit (EC 6.6.1.2)
Aerobic glycerol-3-phosphate dehydrogenase (EC 1.1.5.3)
Alanine dehydrogenase (EC 1.4.1.1)
Alanine racemase (EC 5.1.1.1)

Acetylornithine deacetylase (EC 3.5.1.16)
Acetyltransferase AcuA, acetyl-CoA synthetase inhibitor
Aconitate hydratase (EC 4.2.1.3)
Acyl carrier protein
Acyl-phosphate:glycerol-3-phosphate O-acyltransferase PlsY
Acylphosphate phosphohydrolase (EC 3.6.1.7), putative
Adenine deaminase (EC 3.5.4.2)
Adenine phosphoribosyltransferase (EC 2.4.2.7)
Adenosine deaminase (EC 3.5.4.4)
Adenosylcobinamide-phosphate guanylyltransferase (EC 2.7.7.62)
Adenosylhomocysteinase (EC 3.3.1.1)
Adenosylmethionine-8-amino-7-oxononanoate aminotransferase (EC 2.6.1.62)
Adenylate cyclase (EC 4.6.1.1)
Adenylate kinase (EC 2.7.4.3)
Adenylosuccinate lyase (EC 4.3.2.2)
Adenylosuccinate synthetase (EC 6.3.4.4)
Adenylylsulfate kinase (EC 2.7.1.25)
ADP-ribose pyrophosphatase (EC 3.6.1.13)
Aerobic cobaltochelatase CobS subunit (EC 6.6.1.2)
Aerobic glycerol-3-phosphate dehydrogenase (EC 1.1.5.3)
Alanine dehydrogenase (EC 1.4.1.1)
Alanine racemase (EC 5.1.1.1)

Anthranilate phosphoribosyltransferase (EC 2.4.2.18)
Anthranilate synthase, amidotransferase component (EC 4.1.3.27)
Anthranilate synthase, aminase component (EC 4.1.3.27)
archaeal ATPase, fused to C-terminal DUF234 domain
Archaeal DNA polymerase I (EC 2.7.7.7)
Archaeal DNA polymerase II large subunit (EC 2.7.7.7)
Archaeal DNA polymerase II small subunit (EC 2.7.7.7)
Archaeal S-adenosylmethionine synthetase (EC 2.5.1.6)
Archaeal transcription factor E
archaeosine tRNA-ribosyltransferase (EC 2.4.2.-) type 1
archaeosine tRNA-ribosyltransferase (EC 2.4.2.-) type 2
Archease
Argininosuccinate synthase (EC 6.3.4.5)
Arginyl-tRNA synthetase (EC 6.1.1.19)
Arrested fork binding
Arsenate reductase (EC 1.20.4.1)
Arsenical pump-driving ATPase (EC 3.6.3.16)
Arsenical-resistance protein ACR3
Aspartate aminotransferase (EC 2.6.1.1)
Aspartate carbamoyltransferase (EC 2.1.3.2)
Aspartate-semialdehyde dehydrogenase (EC 1.2.1.11)
Aspartokinase (EC 2.7.2.4)

archaeosine tRNA-ribosyltransferase (EC 2.4.2.-) type 5	Alanyl-tRNA synthetase (EC 6.1.1.7)	Alanyl-tRNA synthetase (EC 6.1.1.7)	Aspartyl-tRNA synthetase (EC 6.1.1.12)
Archaeosortase A	Alcohol dehydrogenase (EC 1.1.1.1)	Alcohol dehydrogenase (EC 1.1.1.1)	Aspartyl-tRNA(Asn) amidotransferase subunit A (EC 6.3.5.6)
Arginine deiminase (EC 3.5.3.6)	Aldehyde dehydrogenase (EC 1.2.1.3)	Aldehyde dehydrogenase (EC 1.2.1.3)	Aspartyl-tRNA(Asn) amidotransferase subunit B (EC 6.3.5.6)
Argininosuccinate lyase (EC 4.3.2.1)	Aldehyde dehydrogenase B (EC 1.2.1.22)	Aldehyde dehydrogenase B (EC 1.2.1.22)	Aspartyl-tRNA(Asn) amidotransferase subunit C (EC 6.3.5.6)
Argininosuccinate synthase (EC 6.3.4.5)	Aldose 1-epimerase (EC 5.1.3.3)	Aldose 1-epimerase (EC 5.1.3.3)	ATP phosphoribosyltransferase (EC 2.4.2.17)
Arginyl-tRNA synthetase (EC 6.1.1.19)	Alkaline phosphatase (EC 3.1.3.1)	Alkaline phosphatase (EC 3.1.3.1)	ATP synthase beta chain (EC 3.6.3.14)
Arrested fork binding	Alkaline phosphatase synthesis transcriptional regulatory protein PhoP	Alkaline phosphatase synthesis transcriptional regulatory protein PhoP	ATP synthase gamma chain (EC 3.6.3.14)
Arsenate reductase (EC 1.20.4.1)	Alkyl hydroperoxide reductase protein F (EC 1.6.4.-)	Alkyl hydroperoxide reductase protein F (EC 1.6.4.-)	ATP-dependent Clp protease ATP-binding subunit ClpA
Arsenical pump-driving ATPase (EC 3.6.3.16)	Alkyl hydroperoxide reductase subunit C-like protein	Alkyl hydroperoxide reductase subunit C-like protein	ATP-dependent DNA helicase pcrA (EC 3.6.1.-)
Arsenical-resistance protein ACR3	Alpha-1,2-mannosidase	Alpha-1,2-mannosidase	ATP-dependent DNA helicase RecG (EC 3.6.1.-)
Arylsulfatase (EC 3.1.6.1)	Alpha-arabinosides ABC transport system, permease protein 1	Alpha-arabinosides ABC transport system, permease protein 1	ATP-dependent DNA helicase UvrD/PcrA
Asparaginyl-tRNA synthetase (EC 6.1.1.22)	Alpha-galactosidase (EC 3.2.1.22)	Alpha-galactosidase (EC 3.2.1.22)	ATP-dependent DNA helicase UvrD/PcrA, proteobacterial paralog
Aspartate aminotransferase (EC 2.6.1.1)	Alpha-glucosidase (EC 3.2.1.20)	Alpha-glucosidase (EC 3.2.1.20)	ATP-dependent DNA ligase (EC 6.5.1.1)
Aspartate carbamoyltransferase (EC 2.1.3.2)	Alpha-N-arabinofuranosidase (EC 3.2.1.55)	Alpha-N-arabinofuranosidase (EC 3.2.1.55)	ATP-dependent hsl protease ATP-binding subunit HslU
Aspartate racemase (EC 5.1.1.13)	Amidophosphoribosyltransferase (EC 2.4.2.14)	Amidophosphoribosyltransferase (EC 2.4.2.14)	ATP-dependent protease La (EC 3.4.21.53)
Aspartate-semialdehyde dehydrogenase (EC 1.2.1.11)	Aminodeoxychorismate lyase (EC 4.1.3.38)	Aminodeoxychorismate lyase (EC 4.1.3.38)	Biotin carboxyl carrier protein of acetyl-CoA carboxylase
Aspartokinase (EC 2.7.2.4)	Ammonium transporter	Ammonium transporter	Biotin carboxylase of acetyl-CoA carboxylase (EC 6.3.4.14)
Aspartyl-tRNA synthetase (EC 6.1.1.12)	Anaerobic glycerol-3-phosphate dehydrogenase subunit A (EC 1.1.5.3)	Anaerobic glycerol-3-phosphate dehydrogenase subunit A (EC 1.1.5.3)	Biotin synthase (EC 2.8.1.6)
Aspartyl-tRNA(Asn) amidotransferase subunit B (EC 6.3.5.6)	Anaerobic glycerol-3-phosphate dehydrogenase subunit B (EC 1.1.5.3)	Anaerobic glycerol-3-phosphate dehydrogenase subunit B (EC 1.1.5.3)	Branched-chain alpha-keto acid dehydrogenase, E1 component, alpha subunit (EC 1.2.4.4)
Aspartyl-tRNA(Asn) amidotransferase subunit C (EC 6.3.5.6)	Anaerobic glycerol-3-phosphate dehydrogenase subunit C (EC 1.1.5.3)	Anaerobic glycerol-3-phosphate dehydrogenase subunit C (EC 1.1.5.3)	Branched-chain alpha-keto acid dehydrogenase, E1 component, beta subunit (EC 1.2.4.4)
ATP phosphoribosyltransferase (EC 2.4.2.17)	Anthranilate phosphoribosyltransferase (EC 2.4.2.18)	Anthranilate phosphoribosyltransferase (EC 2.4.2.18)	Branched-chain amino acid aminotransferase (EC 2.6.1.42)
ATP synthase alpha chain (EC 3.6.3.14)	Anthranilate synthase, amidotransferase component (EC 4.1.3.27)	Anthranilate synthase, amidotransferase component (EC 4.1.3.27)	Branched-chain amino acid transport ATP-binding protein LivG (TC 3.A.1.4.1)
ATP synthase beta chain (EC 3.6.3.14)	Arabinan endo-1,5-alpha-L-arabinosidase (EC 3.2.1.99)	Arabinan endo-1,5-alpha-L-arabinosidase (EC 3.2.1.99)	Butyryl-CoA dehydrogenase (EC 1.3.99.2)

ATP synthase epsilon chain (EC 3.6.3.14)
ATP synthase gamma chain (EC 3.6.3.14)
ATP-dependent Clp protease ATP-binding subunit ClpA
ATP-dependent Clp protease proteolytic subunit (EC 3.4.21.92)
ATP-dependent Clp protease, ATP-binding subunit ClpC
ATP-dependent DNA helicase pcrA (EC 3.6.1.-)
ATP-dependent DNA helicase RecG (EC 3.6.1.-)
ATP-dependent DNA helicase RecQ
ATP-dependent DNA helicase UvrD/PcrA
ATP-dependent DNA helicase UvrD/PcrA, proteobacterial paralog
ATP-dependent DNA helicase, RecQ family
ATP-dependent hsl protease ATP-binding subunit HslU
ATP-dependent protease La (EC 3.4.21.53)
ATP-dependent protease La (EC 3.4.21.53) Type I
ATP-dependent protease La (EC 3.4.21.53) Type II
Beta-carotene ketolase (EC 1.14.-.-)
Beta-hexosaminidase (EC 3.2.1.52)
Betaine--homocysteine S-methyltransferase (EC 2.1.1.5)
Biopolymer transport protein ExbD/TolR
Biotin carboxyl carrier protein of acetyl-CoA carboxylase
Biotin carboxylase of acetyl-CoA carboxylase (EC 6.3.4.14)
Biotin synthase (EC 2.8.1.6)

archaeal ATPase, fused to C-terminal DUF234 domain
Archaeal DNA polymerase I (EC 2.7.7.7)
Archaeal DNA polymerase II large subunit (EC 2.7.7.7)
Archaeal S-adenosylmethionine synthetase (EC 2.5.1.6)
Archaeal transcription factor E
archaeosine tRNA-ribosyltransferase (EC 2.4.2.-) type 1
archaeosine tRNA-ribosyltransferase (EC 2.4.2.-) type 2
archaeosine tRNA-ribosyltransferase (EC 2.4.2.-) type 5
Archease
Arginase (EC 3.5.3.1)
Arginine decarboxylase (EC 4.1.1.19)
Arginine deiminase (EC 3.5.3.6)
Arginine pathway regulatory protein ArgR, repressor of arg regulon
Arginine utilization regulatory protein RocR
Argininosuccinate synthase (EC 6.3.4.5)
Arginyl-tRNA synthetase (EC 6.1.1.19)
Arrested fork binding
Arsenate reductase (EC 1.20.4.1)
Arsenic efflux pump protein
Arsenical pump-driving ATPase (EC 3.6.3.16)
Arsenical resistance operon repressor
Arsenical-resistance protein ACR3

archaeal ATPase, fused to C-terminal DUF234 domain
Archaeal DNA polymerase I (EC 2.7.7.7)
Archaeal DNA polymerase II large subunit (EC 2.7.7.7)
Archaeal S-adenosylmethionine synthetase (EC 2.5.1.6)
Archaeal transcription factor E
archaeosine tRNA-ribosyltransferase (EC 2.4.2.-) type 1
archaeosine tRNA-ribosyltransferase (EC 2.4.2.-) type 2
archaeosine tRNA-ribosyltransferase (EC 2.4.2.-) type 5
Archease
Arginase (EC 3.5.3.1)
Arginine decarboxylase (EC 4.1.1.19)
Arginine deiminase (EC 3.5.3.6)
Arginine pathway regulatory protein ArgR, repressor of arg regulon
Arginine utilization regulatory protein RocR
Argininosuccinate synthase (EC 6.3.4.5)
Arginyl-tRNA synthetase (EC 6.1.1.19)
Arrested fork binding
Arsenate reductase (EC 1.20.4.1)
Arsenic efflux pump protein
Arsenical pump-driving ATPase (EC 3.6.3.16)
Arsenical resistance operon repressor
Arsenical-resistance protein ACR3

C-terminal domain of CinA type S
cAMP-binding proteins - catabolite gene activator and regulatory subunit of cAMP-dependent protein kinases
Candidate Zn-dependent transcriptional regulator, MntR/DtxR family
Carbamoyl-phosphate synthase large chain (EC 6.3.5.5)
Carbamoyl-phosphate synthase small chain (EC 6.3.5.5)
Catalase (EC 1.11.1.6)
Cell division protein FtsA
Cell division protein FtsI [Peptidoglycan synthetase] (EC 2.4.1.129)
Cell division protein FtsZ (EC 3.4.24.-)
Chaperone protein DnaK
Chaperone protein HtpG
ChlD component of cobalt chelatase involved in B12 biosynthesis
Chromosomal replication initiator protein DnaA
Chromosome (plasmid) partitioning protein ParA
Chromosome (plasmid) partitioning protein ParB
Chromosome partition protein smc
Citrate synthase (si) (EC 2.3.3.1)
ClpB protein
Cob(I)alamin adenosyltransferase (EC 2.5.1.17)
Cob(I)alamin adenosyltransferase PduO (EC 2.5.1.17)
Cobalamin biosynthesis protein CbiG
Cobalt-precorrin-3b C17-methyltransferase

Branched-chain alpha-keto acid dehydrogenase, E1 component, alpha subunit (EC 1.2.4.4)	Arylsulfatase (EC 3.1.6.1)	Arylsulfatase (EC 3.1.6.1)	Cobalt-precorrin-6y C5-methyltransferase (EC 2.1.1.-)
Branched-chain amino acid aminotransferase (EC 2.6.1.42)	Asparagine synthetase [glutamine-hydrolyzing] (EC 6.3.5.4) YisO	Asparagine synthetase [glutamine-hydrolyzing] (EC 6.3.5.4) YisO	Cobalt-zinc-cadmium resistance protein
Branched-chain amino acid transport ATP-binding protein LivF (TC 3.A.1.4.1)	Asparaginyl-tRNA synthetase (EC 6.1.1.22)	Asparaginyl-tRNA synthetase (EC 6.1.1.22)	Cobalt-zinc-cadmium resistance protein CzcA
Butyryl-CoA dehydrogenase (EC 1.3.99.2)	Aspartate aminotransferase (EC 2.6.1.1)	Aspartate aminotransferase (EC 2.6.1.1)	CobN component of cobalt chelatase involved in B12 biosynthesis
C-terminal domain of CinA type S	Aspartate ammonia-lyase (EC 4.3.1.1)	Aspartate ammonia-lyase (EC 4.3.1.1)	Cobyrinic acid A,C-diamide synthase
Carbamoyl-phosphate synthase large chain (EC 6.3.5.5)	Aspartate carbamoyltransferase (EC 2.1.3.2)	Aspartate carbamoyltransferase (EC 2.1.3.2)	Coenzyme B synthesis from 2-oxoglutarate: steps 5, 9, and 13
Carbamoyl-phosphate synthase small chain (EC 6.3.5.5)	Aspartate racemase (EC 5.1.1.13)	Aspartate racemase (EC 5.1.1.13)	Coenzyme F420-1:L-glutamate ligase
Catalase (EC 1.11.1.6)	Aspartate-semialdehyde dehydrogenase (EC 1.2.1.11)	Aspartate-semialdehyde dehydrogenase (EC 1.2.1.11)	COG2047: Uncharacterized protein (ATP-grasp superfamily)
CDP-diacylglycerol--glycerol-3-phosphate 3-phosphatidyltransferase (EC 2.7.8.5)	Aspartokinase (EC 2.7.2.4)	Aspartokinase (EC 2.7.2.4)	COG2106 Methylase
Cell division protein FtsA	Aspartyl-tRNA synthetase (EC 6.1.1.12)	Aspartyl-tRNA synthetase (EC 6.1.1.12)	COG3269, Predicted RNA-binding protein, contains TRAM domain
Cell division protein FtsI [Peptidoglycan synthetase] (EC 2.4.1.129)	Aspartyl-tRNA(Asn) amidotransferase subunit A (EC 6.3.5.6)	Aspartyl-tRNA(Asn) amidotransferase subunit A (EC 6.3.5.6)	Copper tolerance protein
Cell division protein FtsX	Aspartyl-tRNA(Asn) amidotransferase subunit B (EC 6.3.5.6)	Aspartyl-tRNA(Asn) amidotransferase subunit B (EC 6.3.5.6)	Copper-containing nitrite reductase (EC 1.7.2.1)
Cell division protein FtsZ (EC 3.4.24.-)	Aspartyl-tRNA(Asn) amidotransferase subunit C (EC 6.3.5.6)	Aspartyl-tRNA(Asn) amidotransferase subunit C (EC 6.3.5.6)	Copper-translocating P-type ATPase (EC 3.6.3.4)
Chaperone protein DnaJ	ATP phosphoribosyltransferase (EC 2.4.2.17)	ATP phosphoribosyltransferase (EC 2.4.2.17)	Coproporphyrinogen III oxidase, oxygen-independent (EC 1.3.99.22)
Chaperone protein DnaK	ATP phosphoribosyltransferase regulatory subunit (EC 2.4.2.17)	ATP phosphoribosyltransferase regulatory subunit (EC 2.4.2.17)	Creatinine amidohydrolase (EC 3.5.2.10)
Chemotaxis protein CheC -- inhibitor of MCP methylation	ATP synthase alpha chain (EC 3.6.3.14)	ATP synthase alpha chain (EC 3.6.3.14)	CRISPR-associated protein Cas1
Chemotaxis protein methyltransferase CheR (EC 2.1.1.80)	ATP synthase beta chain (EC 3.6.3.14)	ATP synthase beta chain (EC 3.6.3.14)	CRISPR-associated protein, Cas5h family
Chorismate synthase (EC 4.2.3.5)	ATP synthase F0 sector subunit a	ATP synthase F0 sector subunit a	CRISPR-associated protein, Csd2/Csh2 family
Chromate transport protein ChrA	ATP synthase gamma chain (EC 3.6.3.14)	ATP synthase gamma chain (EC 3.6.3.14)	Cryptochrome
Chromosomal replication initiator protein DnaA	ATP synthase protein I	ATP synthase protein I	CTP synthase (EC 6.3.4.2)
Chromosome (plasmid) partitioning protein ParA	ATP-dependent Clp protease ATP-binding subunit ClpA	ATP-dependent Clp protease ATP-binding subunit ClpA	Cystathionine beta-synthase (EC 4.2.1.22)

Chromosome partition protein smc	ATP-dependent Clp protease ATP-binding subunit ClpX	ATP-dependent Clp protease ATP-binding subunit ClpX	Cystathionine gamma-lyase (EC 4.4.1.1)
Citrate lyase beta chain (EC 4.1.3.6)	ATP-dependent Clp protease proteolytic subunit (EC 3.4.21.92)	ATP-dependent Clp protease proteolytic subunit (EC 3.4.21.92)	Cystathionine gamma-synthase (EC 2.5.1.48)
Citrate synthase (si) (EC 2.3.3.1)	ATP-dependent Clp protease, ATP-binding subunit ClpC	ATP-dependent Clp protease, ATP-binding subunit ClpC	Cysteine desulfurase (EC 2.8.1.7)
ClpB protein	ATP-dependent DNA helicase pcrA (EC 3.6.1.-)	ATP-dependent DNA helicase pcrA (EC 3.6.1.-)	Cysteine desulfurase (EC 2.8.1.7), SufS subfamily
Cob(I)alamin adenosyltransferase PduO (EC 2.5.1.17)	ATP-dependent DNA helicase RecG (EC 3.6.1.-)	ATP-dependent DNA helicase RecG (EC 3.6.1.-)	Cysteine synthase (EC 2.5.1.47)
Cobalt-precorrin-3b C17-methyltransferase	ATP-dependent DNA helicase RecQ	ATP-dependent DNA helicase RecQ	Cytochrome c heme lyase subunit CcmF
Cobalt-precorrin-4 C11-methyltransferase (EC 2.1.1.133)	ATP-dependent DNA helicase UvrD/PcrA	ATP-dependent DNA helicase UvrD/PcrA	Cytochrome c oxidase (B(O/a)3-type) chain I (EC 1.9.3.1)
Cobalt-zinc-cadmium resistance protein CzcA	ATP-dependent DNA helicase UvrD/PcrA, proteobacterial paralog	ATP-dependent DNA helicase UvrD/PcrA, proteobacterial paralog	Cytochrome c oxidase (B(O/a)3-type) chain II (EC 1.9.3.1)
Coenzyme F420-1:L-glutamate ligase	ATP-dependent DNA helicase, RecQ family	ATP-dependent DNA helicase, RecQ family	Cytochrome c oxidase polypeptide I (EC 1.9.3.1)
COG0488: ATPase components of ABC transporters with duplicated ATPase domains	ATP-dependent DNA ligase (EC 6.5.1.1)	ATP-dependent DNA ligase (EC 6.5.1.1)	Cytochrome c oxidase polypeptide II (EC 1.9.3.1)
COG2106 Methylase	ATP-dependent hsl protease ATP-binding subunit HslU	ATP-dependent hsl protease ATP-binding subunit HslU	Cytochrome c oxidase polypeptide III (EC 1.9.3.1)
COG3269, Predicted RNA-binding protein, contains TRAM domain	ATP-dependent nuclease, subunit A	ATP-dependent nuclease, subunit A	Cytoplasmic axial filament protein CafA and Ribonuclease G (EC 3.1.4.-)
CopG protein	ATP-dependent nuclease, subunit B	ATP-dependent nuclease, subunit B	D-3-phosphoglycerate dehydrogenase (EC 1.1.1.95)
Copper-translocating P-type ATPase (EC 3.6.3.4)	ATP-dependent protease La (EC 3.4.21.53)	ATP-dependent protease La (EC 3.4.21.53)	D-beta-hydroxybutyrate dehydrogenase (EC 1.1.1.30)
Coproporphyrinogen III oxidase, oxygen-independent (EC 1.3.99.22)	ATP-dependent protease La (EC 3.4.21.53) LonB Type I	ATP-dependent protease La (EC 3.4.21.53) LonB Type I	D-glycerate 2-kinase (EC 2.7.1.-)
Creatinine amidohydrolase (EC 3.5.2.10)	ATP-dependent protease La (EC 3.4.21.53) Type I	ATP-dependent protease La (EC 3.4.21.53) Type I	Deacetylases, including yeast histone deacetylase and acetoin utilization protein
CRISPR-associated protein Cas1	ATP-dependent RNA helicase BA2475	ATP-dependent RNA helicase BA2475	Delta-1-pyrroline-5-carboxylate dehydrogenase (EC 1.5.1.12)
CTP synthase (EC 6.3.4.2)	ATP-dependent RNA helicase YfmL	ATP-dependent RNA helicase YfmL	Deoxyadenosine kinase (EC 2.7.1.76)
Cyanobacteria-specific RpoD-like sigma factor, type-1	ATP-dependent RNA helicase YxiN	ATP-dependent RNA helicase YxiN	Deoxyhypusine synthase (EC 2.5.1.46)
Cystathionine beta-synthase (EC 4.2.1.22)	ATPase component BioM of energizing module of biotin ECF transporter	ATPase component BioM of energizing module of biotin ECF transporter	Deoxyribodipyrimidine photolyase (EC 4.1.99.3)
Cystathionine gamma-lyase (EC 4.4.1.1)	Bacteriorhodopsin related protein Blp	Bacteriorhodopsin related protein Blp	Deoxyuridine 5'-triphosphate nucleotidohydrolase (EC 3.6.1.23)
Cysteine desulfurase (EC 2.8.1.7)	Beta-carotene ketolase (EC 1.14.-.-)	Beta-carotene ketolase (EC 1.14.-.-)	Dephospho-CoA kinase archaeal, predicted (EC 2.7.1.24)
Cysteine desulfurase (EC 2.8.1.7), SufS subfamily	Beta-hexosaminidase (EC 3.2.1.52)	Beta-hexosaminidase (EC 3.2.1.52)	Digeranylgeranylglycerophospholipid reductase
Cysteine synthase (EC 2.5.1.47)	Betaine--homocysteine S-methyltransferase (EC 2.1.1.5)	Betaine--homocysteine S-methyltransferase (EC 2.1.1.5)	Dihydrolipoamide acetyltransferase component (E2) of acetoin

			dehydrogenase complex (EC 2.3.1.-)
Cysteinyl-tRNA synthetase (EC 6.1.1.16)	Biopolymer transport protein ExbD/TolR	Biopolymer transport protein ExbD/TolR	Dihydrolipoamide dehydrogenase (EC 1.8.1.4)
Cytidylate kinase (EC 2.7.4.25)	Biosynthetic Aromatic amino acid aminotransferase beta (EC 2.6.1.57)	Biosynthetic Aromatic amino acid aminotransferase beta (EC 2.6.1.57)	Dihydrolipoamide succinyltransferase component (E2) of 2-oxoglutarate dehydrogenase complex (EC 2.3.1.61)
Cytochrome b6-f complex subunit IV (PetD)	Biotin carboxyl carrier protein of acetyl-CoA carboxylase	Biotin carboxyl carrier protein of acetyl-CoA carboxylase	Dihydroxy-acid dehydratase (EC 4.2.1.9)
Cytochrome c heme lyase subunit CcmF	Biotin carboxylase of acetyl-CoA carboxylase (EC 6.3.4.14)	Biotin carboxylase of acetyl-CoA carboxylase (EC 6.3.4.14)	Dimethylallyltransferase (EC 2.5.1.1)
Cytochrome c oxidase (B(O/a)3-type) chain I (EC 1.9.3.1)	Biotin synthase (EC 2.8.1.6)	Biotin synthase (EC 2.8.1.6)	Dipeptide transport ATP-binding protein DppF (TC 3.A.1.5.2)
Cytochrome c oxidase (B(O/a)3-type) chain II (EC 1.9.3.1)	Branched-chain alpha-keto acid dehydrogenase, E1 component, alpha subunit (EC 1.2.4.4)	Branched-chain alpha-keto acid dehydrogenase, E1 component, alpha subunit (EC 1.2.4.4)	Dipeptide transport system permease protein DppB (TC 3.A.1.5.2)
Cytochrome c oxidase polypeptide I (EC 1.9.3.1)	Branched-chain alpha-keto acid dehydrogenase, E1 component, beta subunit (EC 1.2.4.4)	Branched-chain alpha-keto acid dehydrogenase, E1 component, beta subunit (EC 1.2.4.4)	Diphosphomevalonate decarboxylase (EC 4.1.1.33)
Cytochrome c oxidase polypeptide II (EC 1.9.3.1)	Branched-chain amino acid aminotransferase (EC 2.6.1.42)	Branched-chain amino acid aminotransferase (EC 2.6.1.42)	DNA gyrase subunit A (EC 5.99.1.3)
Cytochrome c oxidase polypeptide III (EC 1.9.3.1)	Branched-chain amino acid transport ATP-binding protein LivF (TC 3.A.1.4.1)	Branched-chain amino acid transport ATP-binding protein LivF (TC 3.A.1.4.1)	DNA gyrase subunit B (EC 5.99.1.3)
Cytochrome c-type biogenesis protein CcmC, putative heme lyase for CcmE	Branched-chain amino acid transport ATP-binding protein LivG (TC 3.A.1.4.1)	Branched-chain amino acid transport ATP-binding protein LivG (TC 3.A.1.4.1)	DNA helicase (Rad25 homolog)
Cytochrome c-type biogenesis protein CcmE, heme chaperone	Branched-chain amino acid transport system permease protein LivM (TC 3.A.1.4.1)	Branched-chain amino acid transport system permease protein LivM (TC 3.A.1.4.1)	DNA ligase (EC 6.5.1.2)
Cytochrome c551 peroxidase (EC 1.11.1.5)	Brp-like protein	Brp-like protein	DNA mismatch repair protein MutL
Cytochrome d ubiquinol oxidase subunit I (EC 1.10.3.-)	Butyryl-CoA dehydrogenase (EC 1.3.99.2)	Butyryl-CoA dehydrogenase (EC 1.3.99.2)	DNA mismatch repair protein MutS
Cytoplasmic axial filament protein CafA and Ribonuclease G (EC 3.1.4.-)	C-terminal domain of CinA type S	C-terminal domain of CinA type S	DNA polymerase III alpha subunit (EC 2.7.7.7)
D-3-phosphoglycerate dehydrogenase (EC 1.1.1.95)	Cadmium efflux system accessory protein	Cadmium efflux system accessory protein	DNA polymerase III delta prime subunit (EC 2.7.7.7)
D-alanine aminotransferase (EC 2.6.1.21)	Cadmium-transporting ATPase (EC 3.6.3.3)	Cadmium-transporting ATPase (EC 3.6.3.3)	DNA polymerase III delta subunit (EC 2.7.7.7)
D-alanine--D-alanine ligase (EC 6.3.2.4)	Carbamoyl-phosphate synthase large chain (EC 6.3.5.5)	Carbamoyl-phosphate synthase large chain (EC 6.3.5.5)	DNA polymerase III subunits gamma and tau (EC 2.7.7.7)
D-alanyl-D-alanine carboxypeptidase (EC 3.4.16.4)	Carbamoyl-phosphate synthase small chain (EC 6.3.5.5)	Carbamoyl-phosphate synthase small chain (EC 6.3.5.5)	DNA polymerase IV (EC 2.7.7.7)
D-amino acid dehydrogenase small subunit (EC 1.4.99.1)	Carbon starvation protein A	Carbon starvation protein A	DNA polymerase sliding clamp protein PCNA
D-glycerate 3-kinase (EC 2.7.1.31), plant type	Carbonic anhydrase (EC 4.2.1.1)	Carbonic anhydrase (EC 4.2.1.1)	DNA primase large subunit (EC 2.7.7.-)
D-hydantoinase (EC 3.5.2.2)	Cardiolipin synthetase (EC 2.7.8.-)	Cardiolipin synthetase (EC 2.7.8.-)	DNA primase small subunit (EC 2.7.7.-)
D-mannose isomerase (EC 5.3.1.7)	Catabolite control protein A	Catabolite control protein A	DNA repair and recombination protein RadA

Death on curing protein, Doc toxin	Catalase (EC 1.11.1.6)	Catalase (EC 1.11.1.6)	DNA repair and recombination protein RadB
Delta-1-pyrroline-5-carboxylate dehydrogenase (EC 1.5.1.12)	Catalyzes the cleavage of p-aminobenzoyl-glutamate to p-aminobenzoate and glutamate, subunit A	Catalyzes the cleavage of p-aminobenzoyl-glutamate to p-aminobenzoate and glutamate, subunit A	DNA repair exonuclease family protein YhaO
Deoxyadenosine kinase (EC 2.7.1.76)	CDP-diacylglycerol--glycerol-3-phosphate 3-phosphatidyltransferase (EC 2.7.8.5)	CDP-diacylglycerol--glycerol-3-phosphate 3-phosphatidyltransferase (EC 2.7.8.5)	DNA repair protein RecN
Deoxycytidine triphosphate deaminase (EC 3.5.4.13)	CDP-diacylglycerol--serine O-phosphatidyltransferase (EC 2.7.8.8)	CDP-diacylglycerol--serine O-phosphatidyltransferase (EC 2.7.8.8)	DNA replication helicase protein MCM
Deoxyguanosinetriphosphate triphosphohydrolase (EC 3.1.5.1)	Cell division protein FtsA	Cell division protein FtsA	DNA topoisomerase I (EC 5.99.1.2)
Deoxyhypusine synthase (EC 2.5.1.46)	Cell division protein FtsI [Peptidoglycan synthetase] (EC 2.4.1.129)	Cell division protein FtsI [Peptidoglycan synthetase] (EC 2.4.1.129)	DNA topoisomerase I (EC 5.99.1.2)
Deoxyribodipyrimidine photolyase (EC 4.1.99.3)	Cell division protein FtsK	Cell division protein FtsK	DNA topoisomerase VI subunit B (EC 5.99.1.3)
Deoxyribodipyrimidine photolyase, type II (EC 4.1.99.3)	Cell division protein FtsW	Cell division protein FtsW	DNA-directed RNA polymerase alpha subunit (EC 2.7.7.6)
Deoxyribose-phosphate aldolase (EC 4.1.2.4)	Cell division protein FtsX	Cell division protein FtsX	DNA-directed RNA polymerase beta' subunit (EC 2.7.7.6)
Deoxyuridine 5'-triphosphate nucleotidohydrolase (EC 3.6.1.23)	Cell division protein FtsZ (EC 3.4.24.-)	Cell division protein FtsZ (EC 3.4.24.-)	DNA-directed RNA polymerase subunit A' (EC 2.7.7.6)
Diaminopimelate decarboxylase (EC 4.1.1.20)	Cell-shape determining protein MreBH	Cell-shape determining protein MreBH	DNA-directed RNA polymerase subunit A'' (EC 2.7.7.6)
Digeranylgeranylglycerophospholipid reductase	Chaperone protein DnaJ	Chaperone protein DnaJ	DNA-directed RNA polymerase subunit B' (EC 2.7.7.6)
Dihydrolipoamide acetyltransferase component of pyruvate dehydrogenase complex (EC 2.3.1.12)	Chaperone protein DnaK	Chaperone protein DnaK	DNA-directed RNA polymerase subunit B'' (EC 2.7.7.6)
Dihydrolipoamide dehydrogenase (EC 1.8.1.4)	Chemotaxis protein CheC -- inhibitor of MCP methylation	Chemotaxis protein CheC -- inhibitor of MCP methylation	DNA-directed RNA polymerase subunit D (EC 2.7.7.6)
Dihydrolipoamide dehydrogenase of 2-oxoglutarate dehydrogenase (EC 1.8.1.4)	Chemotaxis protein CheV (EC 2.7.3.-)	Chemotaxis protein CheV (EC 2.7.3.-)	DNA-directed RNA polymerase subunit E' (EC 2.7.7.6)
Dihydrolipoamide succinyltransferase component (E2) of 2-oxoglutarate dehydrogenase complex (EC 2.3.1.61)	Chemotaxis protein methyltransferase CheR (EC 2.1.1.80)	Chemotaxis protein methyltransferase CheR (EC 2.1.1.80)	dTDP-4-dehydrorhamnose 3,5-epimerase (EC 5.1.3.13)
Dihydroneopterin aldolase (EC 4.1.2.25)	Chemotaxis response regulator protein-glutamate methylesterase CheB (EC 3.1.1.61)	Chemotaxis response regulator protein-glutamate methylesterase CheB (EC 3.1.1.61)	dTDP-glucose 4,6-dehydratase (EC 4.2.1.46)
Dihydroorotase (EC 3.5.2.3)	Chorismate synthase (EC 4.2.3.5)	Chorismate synthase (EC 4.2.3.5)	DUF124 domain-containing protein
Dihydroorotate dehydrogenase (EC 1.3.3.1)	Chromosomal replication initiator protein DnaA	Chromosomal replication initiator protein DnaA	Electron transfer flavoprotein, alpha subunit
Dihydroorotate dehydrogenase, catalytic subunit (EC 1.3.3.1)	Chromosome (plasmid) partitioning protein ParA	Chromosome (plasmid) partitioning protein ParA	Electron transfer flavoprotein, beta subunit

Dihydropteroate synthase (EC 2.5.1.15)	Chromosome (plasmid) partitioning protein ParB	Chromosome (plasmid) partitioning protein ParB	Endo-1,4-beta-xylanase A precursor (EC 3.2.1.8)
Dihydroxy-acid dehydratase (EC 4.2.1.9)	Chromosome (plasmid) partitioning protein ParB-2	Chromosome (plasmid) partitioning protein ParB-2	Endonuclease IV (EC 3.1.21.2)
Dimethylhistidine N-methyltransferase (EC 2.1.1.44)	Chromosome partition protein smc	Chromosome partition protein smc	Enolase (EC 4.2.1.11)
Dipeptide transport system permease protein DppC (TC 3.A.1.5.2)	Chromosome replication initiation protein dnaD	Chromosome replication initiation protein dnaD	Enoyl-[acyl-carrier-protein] reductase [FMN] (EC 1.3.1.9)
Diphosphomevalonate decarboxylase (EC 4.1.1.33)	Citrate synthase (si) (EC 2.3.3.1)	Citrate synthase (si) (EC 2.3.3.1)	Enoyl-[acyl-carrier-protein] reductase [NADH] (EC 1.3.1.9)
Diphthine synthase (EC 2.1.1.98)	ClpB protein	ClpB protein	Enoyl-CoA hydratase (EC 4.2.1.17)
diptheria toxin resistance protein required for diphthamide biosynthesis (Saccharomyces)-like 1	Cob(I)alamin adenosyltransferase (EC 2.5.1.17)	Cob(I)alamin adenosyltransferase (EC 2.5.1.17)	Error-prone, lesion bypass DNA polymerase V (UmuC)
DNA gyrase subunit A (EC 5.99.1.3)	Cob(I)alamin adenosyltransferase PduO (EC 2.5.1.17)	Cob(I)alamin adenosyltransferase PduO (EC 2.5.1.17)	Eukaryotic peptide chain release factor subunit 1
DNA gyrase subunit B (EC 5.99.1.3)	Cobalamin synthase	Cobalamin synthase	Eukaryotic translation initiation factor 2 alpha subunit
DNA helicase (Rad25 homolog)	Cobalt-precorrin-3b C17-methyltransferase	Cobalt-precorrin-3b C17-methyltransferase	Eukaryotic translation initiation factor 2 beta subunit
DNA ligase (EC 6.5.1.2)	Cobalt-precorrin-4 C11-methyltransferase (EC 2.1.1.133)	Cobalt-precorrin-4 C11-methyltransferase (EC 2.1.1.133)	Eukaryotic translation initiation factor 5A
DNA mismatch repair protein MutL	Cobalt-precorrin-8x methylmutase (EC 5.4.1.2)	Cobalt-precorrin-8x methylmutase (EC 5.4.1.2)	Excinuclease ABC subunit A
DNA mismatch repair protein MutS	Cobalt-zinc-cadmium resistance protein	Cobalt-zinc-cadmium resistance protein	Excinuclease ABC subunit A paralog in greater Bacteroides group
DNA polymerase II (EC 2.7.7.7)	Cobalt-zinc-cadmium resistance protein CzcA	Cobalt-zinc-cadmium resistance protein CzcA	Excinuclease ABC subunit B
DNA polymerase III alpha subunit (EC 2.7.7.7)	Cobalt-zinc-cadmium resistance protein CzcD	Cobalt-zinc-cadmium resistance protein CzcD	Excinuclease ABC subunit C
DNA polymerase III beta subunit (EC 2.7.7.7)	CobN component of cobalt chelatase involved in B12 biosynthesis	CobN component of cobalt chelatase involved in B12 biosynthesis	Exonuclease SbcC
DNA polymerase III delta prime subunit (EC 2.7.7.7)	Cobyrinic acid A,C-diamide synthase	Cobyrinic acid A,C-diamide synthase	Fe-S OXIDOREDUCTASE (1.8.-.-) Wyeosine biosynthesis
DNA polymerase III delta subunit (EC 2.7.7.7)	Coenzyme gamma-F420-2:L-glutamate ligase	Coenzyme gamma-F420-2:L-glutamate ligase	Ferredoxin
DNA polymerase III subunits gamma and tau (EC 2.7.7.7)	COG0488: ATPase components of ABC transporters with duplicated ATPase domains	COG0488: ATPase components of ABC transporters with duplicated ATPase domains	Ferredoxin reductase
DNA polymerase IV (EC 2.7.7.7)	COG2047: Uncharacterized protein (ATP-grasp superfamily)	COG2047: Uncharacterized protein (ATP-grasp superfamily)	Ferredoxin--nitrite reductase (EC 1.7.7.1)
DNA polymerase sliding clamp protein PCNA	COG3269, Predicted RNA-binding protein, contains TRAM domain	COG3269, Predicted RNA-binding protein, contains TRAM domain	Ferric iron ABC transporter, permease protein
DNA primase (EC 2.7.-.-)	COG4123: Predicted O-methyltransferase	COG4123: Predicted O-methyltransferase	Ferric siderophore transport system, biopolymer transport protein ExbB
DNA repair and recombination protein RadA	CopG protein	CopG protein	Ferric siderophore transport system, periplasmic binding protein TonB

DNA repair and recombination protein RadB	Copper-containing nitrite reductase (EC 1.7.2.1)	Copper-containing nitrite reductase (EC 1.7.2.1)	Ferric uptake regulation protein FUR
DNA repair exonuclease family protein YhaO	Copper-translocating P-type ATPase (EC 3.6.3.4)	Copper-translocating P-type ATPase (EC 3.6.3.4)	Ferrochelatase, protoheme ferro-lyase (EC 4.99.1.1)
DNA repair protein RadC	Coproporphyrinogen III oxidase, oxygen-independent (EC 1.3.99.22)	Coproporphyrinogen III oxidase, oxygen-independent (EC 1.3.99.22)	FIG016027: protein of unknown function YeaO
DNA replication helicase protein MCM	Creatinine amidohydrolase (EC 3.5.2.10)	Creatinine amidohydrolase (EC 3.5.2.10)	Flagella-related protein FlaE
DNA topoisomerase I (EC 5.99.1.2)	CRISPR-associated helicase Cas3	CRISPR-associated helicase Cas3	Flagella-related protein FlaH
DNA topoisomerase VI subunit A (EC 5.99.1.3)	CRISPR-associated protein Cas1	CRISPR-associated protein Cas1	Flagella-related protein FlaI
DNA-directed RNA polymerase alpha subunit (EC 2.7.7.6)	CRISPR-associated RAMP Cmr6	CRISPR-associated RAMP Cmr6	Flagella-related protein FlaJ
DNA-directed RNA polymerase beta' subunit (EC 2.7.7.6)	CRISPR-associated RecB family exonuclease Cas4a	CRISPR-associated RecB family exonuclease Cas4a	Flagellar basal-body rod protein FlgG
DNA-directed RNA polymerase subunit A' (EC 2.7.7.6)	crotonyl-CoA reductase	crotonyl-CoA reductase	Flagellar biosynthesis protein FlhA
DNA-directed RNA polymerase subunit A" (EC 2.7.7.6)	CsdL (EC-YgdL) protein of the HesA/MoeB/ThiF family, part of the CsdA-E-L sulfur transfer pathway	CsdL (EC-YgdL) protein of the HesA/MoeB/ThiF family, part of the CsdA-E-L sulfur transfer pathway	Flagellar biosynthesis protein FlhB
DNA-directed RNA polymerase subunit B' (EC 2.7.7.6)	CTP synthase (EC 6.3.4.2)	CTP synthase (EC 6.3.4.2)	Flagellar biosynthesis protein FliR
DNA-directed RNA polymerase subunit B" (EC 2.7.7.6)	Cyclic pyranopterin phosphate synthase (MoaA) (EC 4.1.99.18)	Cyclic pyranopterin phosphate synthase (MoaA) (EC 4.1.99.18)	Flagellar hook-associated protein FliD
DNA-directed RNA polymerase subunit D (EC 2.7.7.6)	Cystathionine beta-synthase (EC 4.2.1.22)	Cystathionine beta-synthase (EC 4.2.1.22)	Flagellar synthesis regulator FleN
DNA-directed RNA polymerase subunit E' (EC 2.7.7.6)	Cystathionine gamma-lyase (EC 4.4.1.1)	Cystathionine gamma-lyase (EC 4.4.1.1)	Flagellin FlaA1
DNA-directed RNA polymerase subunit F (EC 2.7.7.6)	Cysteine desulfurase (EC 2.8.1.7)	Cysteine desulfurase (EC 2.8.1.7)	Flagellin protein FlaA
DNA-directed RNA polymerase subunit H (EC 2.7.7.6)	Cysteine desulfurase (EC 2.8.1.7), SufS subfamily	Cysteine desulfurase (EC 2.8.1.7), SufS subfamily	Flagellum-specific ATP synthase FliI
DNA-directed RNA polymerase subunit N (EC 2.7.7.6)	Cysteine synthase (EC 2.5.1.47)	Cysteine synthase (EC 2.5.1.47)	Flap structure-specific endonuclease (EC 3.-.-.-)
dTDP-4-dehydrorhamnose 3,5-epimerase (EC 5.1.3.13)	Cysteinyl-tRNA synthetase (EC 6.1.1.16)	Cysteinyl-tRNA synthetase (EC 6.1.1.16)	Formate--tetrahydrofolate ligase (EC 6.3.4.3)
dTDP-4-dehydrorhamnose reductase (EC 1.1.1.133)	Cytidylate kinase (EC 2.7.4.25)	Cytidylate kinase (EC 2.7.4.25)	Formiminoglutamic iminohydrolase (EC 3.5.3.13)
dTDP-glucose 4,6-dehydratase (EC 4.2.1.46)	Cytochrome b6-f complex subunit, apocytochrome f	Cytochrome b6-f complex subunit, apocytochrome f	Fructose-1,6-bisphosphatase, type I (EC 3.1.3.11)
Electron transfer flavoprotein, alpha subunit	Cytochrome c heme lyase subunit CcmF	Cytochrome c heme lyase subunit CcmF	Fructose-bisphosphate aldolase class I (EC 4.1.2.13)
Electron transfer flavoprotein, beta subunit	Cytochrome c oxidase (B(O/a)3-type) chain I (EC 1.9.3.1)	Cytochrome c oxidase (B(O/a)3-type) chain I (EC 1.9.3.1)	Fructose-bisphosphate aldolase, archaeal class I (EC 4.1.2.13)
Endonuclease III (EC 4.2.99.18)	Cytochrome c oxidase (B(O/a)3-type) chain II (EC 1.9.3.1)	Cytochrome c oxidase (B(O/a)3-type) chain II (EC 1.9.3.1)	Fumarate hydratase class II (EC 4.2.1.2)
Endonuclease IV (EC 3.1.21.2)	Cytochrome c oxidase polypeptide I (EC 1.9.3.1)	Cytochrome c oxidase polypeptide I (EC 1.9.3.1)	Fumarylacetoacetase (EC 3.7.1.2)
Endonuclease V (EC 3.1.21.7)	Cytochrome c oxidase polypeptide II (EC 1.9.3.1)	Cytochrome c oxidase polypeptide II (EC 1.9.3.1)	Galactitol-1-phosphate 5-dehydrogenase (EC 1.1.1.251)

Enolase (EC 4.2.1.11)
Enoyl-[acyl-carrier-protein] reductase [FMN] (EC 1.3.1.9)
Enoyl-[acyl-carrier-protein] reductase [NADH] (EC 1.3.1.9)
Enoyl-CoA hydratase (EC 4.2.1.17)
Error-prone, lesion bypass DNA polymerase V (UmuC)
Eukaryotic peptide chain release factor subunit 1
Eukaryotic translation initiation factor 2 gamma subunit
Excinuclease ABC subunit A
Excinuclease ABC subunit A paralog in greater Bacteroides group
Excinuclease ABC subunit B
Excinuclease ABC subunit C
Exonuclease SbcD
Expressed protein possibly involved in photorespiration
Fe-S OXIDOREDUCTASE (1.8.-.-) Wyeosine biosynthesis
Ferredoxin
Ferredoxin reductase
Ferredoxin--NADP(+) reductase (EC 1.18.1.2)
Ferredoxin--nitrite reductase (EC 1.7.7.1)
Ferric siderophore transport system, biopolymer transport protein ExbB
Ferric siderophore transport system, periplasmic binding protein TonB
Ferrochelatase, protoheme ferro-lyase (EC 4.99.1.1)

Cytochrome c oxidase polypeptide III (EC 1.9.3.1)
Cytochrome c-type biogenesis protein CcdA (DsbD analog)
Cytochrome c-type biogenesis protein CcmC, putative heme lyase for CcmE
Cytochrome c-type biogenesis protein CcmE, heme chaperone
Cytochrome c-type biogenesis protein CcmG/DsbE, thiol:disulfide oxidoreductase
Cytochrome c-type biogenesis protein Ccs1/ResB
Cytochrome c-type biogenesis protein CcsA/ResC
Cytochrome c-type biogenesis protein DsbD, protein-disulfide reductase (EC 1.8.1.8)
Cytochrome d ubiquinol oxidase subunit I (EC 1.10.3.-)
Cytochrome d ubiquinol oxidase subunit II (EC 1.10.3.-)
Cytochrome oxidase biogenesis protein Sco1/SenC/PrrC, putative copper metallochaperone
Cytoplasmic axial filament protein CafA and Ribonuclease G (EC 3.1.4.-)
D-3-phosphoglycerate dehydrogenase (EC 1.1.1.95)
D-alanine aminotransferase (EC 2.6.1.21)
D-alanine--D-alanine ligase (EC 6.3.2.4)
D-alanyl-D-alanine carboxypeptidase (EC 3.4.16.4)
D-amino acid dehydrogenase small subunit (EC 1.4.99.1)
D-arabino-3-hexulose 6-phosphate formaldehyde-lyase (EC 4.1.2.43)
D-beta-hydroxybutyrate dehydrogenase (EC 1.1.1.30)
D-beta-hydroxybutyrate permease
D-Lactate dehydrogenase, cytochrome c-dependent (EC 1.1.2.4)

Cytochrome c oxidase polypeptide III (EC 1.9.3.1)
Cytochrome c-type biogenesis protein CcdA (DsbD analog)
Cytochrome c-type biogenesis protein CcmC, putative heme lyase for CcmE
Cytochrome c-type biogenesis protein CcmE, heme chaperone
Cytochrome c-type biogenesis protein CcmG/DsbE, thiol:disulfide oxidoreductase
Cytochrome c-type biogenesis protein Ccs1/ResB
Cytochrome c-type biogenesis protein CcsA/ResC
Cytochrome c-type biogenesis protein DsbD, protein-disulfide reductase (EC 1.8.1.8)
Cytochrome d ubiquinol oxidase subunit I (EC 1.10.3.-)
Cytochrome d ubiquinol oxidase subunit II (EC 1.10.3.-)
Cytochrome oxidase biogenesis protein Sco1/SenC/PrrC, putative copper metallochaperone
Cytoplasmic axial filament protein CafA and Ribonuclease G (EC 3.1.4.-)
D-3-phosphoglycerate dehydrogenase (EC 1.1.1.95)
D-alanine aminotransferase (EC 2.6.1.21)
D-alanine--D-alanine ligase (EC 6.3.2.4)
D-alanyl-D-alanine carboxypeptidase (EC 3.4.16.4)
D-amino acid dehydrogenase small subunit (EC 1.4.99.1)
D-arabino-3-hexulose 6-phosphate formaldehyde-lyase (EC 4.1.2.43)
D-beta-hydroxybutyrate dehydrogenase (EC 1.1.1.30)
D-beta-hydroxybutyrate permease
D-Lactate dehydrogenase, cytochrome c-dependent (EC 1.1.2.4)

Gamma-glutamyl phosphate reductase (EC 1.2.1.41)
Gamma-glutamyltranspeptidase (EC 2.3.2.2)
gliding motility protein GldF
Glucoamylase (EC 3.2.1.3)
Gluconate dehydratase (EC 4.2.1.39)
Glucosamine--fructose-6-phosphate aminotransferase [isomerizing] (EC 2.6.1.16)
Glucose-1-phosphate thymidylyltransferase (EC 2.7.7.24)
Glucose-6-phosphate 1-dehydrogenase (EC 1.1.1.49)
Glucose-6-phosphate isomerase (EC 5.3.1.9)
Glucose-methanol-choline (GMC) oxidoreductase:NAD binding site
Glutamate synthase [NADPH] large chain (EC 1.4.1.13)
Glutamate synthase [NADPH] small chain (EC 1.4.1.13)
Glutamate--cysteine ligase archaeal (EC 6.3.2.2)
Glutamate-1-semialdehyde aminotransferase (EC 5.4.3.8)
Glutamine synthetase type I (EC 6.3.1.2)
Glutaminyl-tRNA synthetase (EC 6.1.1.18)
Glutamyl-tRNA(Gln) amidotransferase subunit A (EC 6.3.5.7)
Glutamyl-tRNA(Gln) amidotransferase subunit B (EC 6.3.5.7)
Glutamyl-tRNA(Gln) amidotransferase transferase subunit (EC 6.3.5.7)
Glutamyl-tRNA(Gln) synthetase (EC 6.1.1.24)
Glutaredoxin-related protein

Ferrous iron transport protein B	DEAD-box ATP-dependent RNA helicase CshB (EC 3.6.4.13)	DEAD-box ATP-dependent RNA helicase CshB (EC 3.6.4.13)	Glutathione S-transferase, omega (EC 2.5.1.18)
Fibrillarin	Death on curing protein, Doc toxin	Death on curing protein, Doc toxin	Glycerol dehydrogenase (EC 1.1.1.6)
FIG000605: protein co-occurring with transport systems (COG1739)	Deblocking aminopeptidase (EC 3.4.11.-)	Deblocking aminopeptidase (EC 3.4.11.-)	Glycerol kinase (EC 2.7.1.30)
FKBP-type peptidyl-prolyl cis-trans isomerase SlyD (EC 5.2.1.8)	Delta-1-pyrroline-5-carboxylate dehydrogenase (EC 1.5.1.12)	Delta-1-pyrroline-5-carboxylate dehydrogenase (EC 1.5.1.12)	Glycerol-1-phosphate dehydrogenase [NAD(P)] (EC 1.1.1.261)
Flagella-related protein FlaH	Deoxyadenosine kinase (EC 2.7.1.76)	Deoxyadenosine kinase (EC 2.7.1.76)	Glycerol-3-phosphate ABC transporter, ATP-binding protein UgpC (TC 3.A.1.1.3)
Flagella-related protein FlaI	Deoxycytidine triphosphate deaminase (EC 3.5.4.13)	Deoxycytidine triphosphate deaminase (EC 3.5.4.13)	Glycerol-3-phosphate ABC transporter, periplasmic glycerol-3-phosphate-binding protein (TC 3.A.1.1.3)
Flagella-related protein FlaJ	Deoxyguanosine kinase (EC 2.7.1.113)	Deoxyguanosine kinase (EC 2.7.1.113)	Glycerol-3-phosphate ABC transporter, permease protein UgpA (TC 3.A.1.1.3)
Flagellar biosynthesis protein FlhA	Deoxyguanosinetriphosphate triphosphohydrolase (EC 3.1.5.1)	Deoxyguanosinetriphosphate triphosphohydrolase (EC 3.1.5.1)	Glycerol-3-phosphate ABC transporter, permease protein UgpE (TC 3.A.1.1.3)
Flagellar biosynthesis protein FlhB	Deoxyhypusine synthase (EC 2.5.1.46)	Deoxyhypusine synthase (EC 2.5.1.46)	Glycerol-3-phosphate regulon repressor, DeoR family
Flagellar biosynthesis protein FliR	Deoxyribodipyrimidine photolyase (EC 4.1.99.3)	Deoxyribodipyrimidine photolyase (EC 4.1.99.3)	Glycerophosphoryl diester phosphodiesterase (EC 3.1.4.46)
Flagellar hook protein FlgE	Deoxyribodipyrimidine photolyase, type II (EC 4.1.99.3)	Deoxyribodipyrimidine photolyase, type II (EC 4.1.99.3)	Glycine cleavage system H protein
Flagellar hook-associated protein FlgK	Deoxyribose-phosphate aldolase (EC 4.1.2.4)	Deoxyribose-phosphate aldolase (EC 4.1.2.4)	Glycine dehydrogenase [decarboxylating] (glycine cleavage system P protein) (EC 1.4.4.2)
Flagellar hook-associated protein FliD	Deoxyuridine 5'-triphosphate nucleotidohydrolase (EC 3.6.1.23)	Deoxyuridine 5'-triphosphate nucleotidohydrolase (EC 3.6.1.23)	Glycine dehydrogenase [decarboxylating] (glycine cleavage system P2 protein) (EC 1.4.4.2)
Flagellar hook-basal body complex protein FliE	Dephospho-CoA kinase (EC 2.7.1.24)	Dephospho-CoA kinase (EC 2.7.1.24)	Glycogen phosphorylase (EC 2.4.1.1)
Flagellar hook-length control protein FliK	Diaminopimelate decarboxylase (EC 4.1.1.20)	Diaminopimelate decarboxylase (EC 4.1.1.20)	GTP cyclohydrolase I (EC 3.5.4.16) type 1
Flagellar motor rotation protein MotA	Diaminopimelate epimerase (EC 5.1.1.7)	Diaminopimelate epimerase (EC 5.1.1.7)	GTP cyclohydrolase I (EC 3.5.4.16) type 2
Flagellar motor switch protein FliM	Digeranylgeranylglycerophospholipid reductase	Digeranylgeranylglycerophospholipid reductase	GTP pyrophosphokinase (EC 2.7.6.5), (p)ppGpp synthetase II
Flagellin protein FlaA	Dihydrofolate reductase (EC 1.5.1.3)	Dihydrofolate reductase (EC 1.5.1.3)	GTP-binding and nucleic acid-binding protein YchF
Flap structure-specific endonuclease (EC 3.-.-.-)	Dihydrolipoamide acyltransferase component of branched-chain alpha-keto acid dehydrogenase complex (EC 2.3.1.168)	Dihydrolipoamide acyltransferase component of branched-chain alpha-keto acid dehydrogenase complex (EC 2.3.1.168)	GTP-binding protein EngB
Formate dehydrogenase-O, major subunit (EC 1.2.1.2)	Dihydrolipoamide dehydrogenase (EC 1.8.1.4)	Dihydrolipoamide dehydrogenase (EC 1.8.1.4)	GTP-binding protein HflX
Formate--tetrahydrofolate ligase (EC 6.3.4.3)	Dihydrolipoamide dehydrogenase of branched-chain alpha-keto acid dehydrogenase (EC 1.8.1.4)	Dihydrolipoamide dehydrogenase of branched-chain alpha-keto acid dehydrogenase (EC 1.8.1.4)	GTPase and tRNA-U34 5-formylation enzyme TrmE

Formyltetrahydrofolate deformylase (EC 3.5.1.10)	Dihydrolipoamide dehydrogenase of pyruvate dehydrogenase complex (EC 1.8.1.4)	Dihydrolipoamide dehydrogenase of pyruvate dehydrogenase complex (EC 1.8.1.4)	Heat shock protein 60 family chaperone GroEL
Fructose-1,6-bisphosphatase, type I (EC 3.1.3.11)	Dihydrolipoamide succinyltransferase component (E2) of 2-oxoglutarate dehydrogenase complex (EC 2.3.1.61)	Dihydrolipoamide succinyltransferase component (E2) of 2-oxoglutarate dehydrogenase complex (EC 2.3.1.61)	Heat shock protein 60 family co-chaperone GroES
Fructose-bisphosphate aldolase class I (EC 4.1.2.13)	Dihydroorotase (EC 3.5.2.3)	Dihydroorotase (EC 3.5.2.3)	Heme biosynthesis protein related to NirG
Fumarate hydratase class II (EC 4.2.1.2)	Dihydroorotate dehydrogenase (EC 1.3.3.1)	Dihydroorotate dehydrogenase (EC 1.3.3.1)	Heme biosynthesis protein related to NirL
Fumarylacetoacetase (EC 3.7.1.2)	Dihydroorotate dehydrogenase electron transfer subunit (EC 1.3.3.1)	Dihydroorotate dehydrogenase electron transfer subunit (EC 1.3.3.1)	Heme O synthase, protoheme IX farnesyltransferase (EC 2.5.1.-) COX10-CtaB
Galactose-1-phosphate uridylyltransferase (EC 2.7.7.10)	Dihydropteroate synthase (EC 2.5.1.15)	Dihydropteroate synthase (EC 2.5.1.15)	Hexokinase (EC 2.7.1.1)
Gamma-glutamyl phosphate reductase (EC 1.2.1.41)	Dihydropyrimidinase (EC 3.5.2.2)	Dihydropyrimidinase (EC 3.5.2.2)	High-affinity branched-chain amino acid transport system permease protein LivH (TC 3.A.1.4.1)
gliding motility protein GldF	Dihydropyrimidine dehydrogenase [NADP+] (EC 1.3.1.2)	Dihydropyrimidine dehydrogenase [NADP+] (EC 1.3.1.2)	Histidinol dehydrogenase (EC 1.1.1.23)
Glucoamylase (EC 3.2.1.3)	Dihydroxy-acid dehydratase (EC 4.2.1.9)	Dihydroxy-acid dehydratase (EC 4.2.1.9)	Histidinol-phosphatase [alternative form] (EC 3.1.3.15)
Glucokinase (EC 2.7.1.2)	Dihydroxyacetone kinase family protein	Dihydroxyacetone kinase family protein	Histidinol-phosphate aminotransferase (EC 2.6.1.9)
Gluconate dehydratase (EC 4.2.1.39)	Dimeric dUTPase (EC 3.6.1.23)	Dimeric dUTPase (EC 3.6.1.23)	Holliday junction DNA helicase RuvA
Glucosamine--fructose-6-phosphate aminotransferase [isomerizing] (EC 2.6.1.16)	Dimethylallyltransferase (EC 2.5.1.1)	Dimethylallyltransferase (EC 2.5.1.1)	Homogentisate 1,2-dioxygenase (EC 1.13.11.5)
Glucose 1-dehydrogenase (EC 1.1.1.47)	Dimethylhistidine N-methyltransferase (EC 2.1.1.44)	Dimethylhistidine N-methyltransferase (EC 2.1.1.44)	Homoserine dehydrogenase (EC 1.1.1.3)
Glucose-1-phosphate cytidylyltransferase (EC 2.7.7.33)	DinG family ATP-dependent helicase CPE1197	DinG family ATP-dependent helicase CPE1197	Homoserine kinase (EC 2.7.1.39)
Glucose-1-phosphate thymidylyltransferase (EC 2.7.7.24)	DinG family ATP-dependent helicase YoaA	DinG family ATP-dependent helicase YoaA	Homoserine O-acetyltransferase (EC 2.3.1.31)
Glucose-6-phosphate 1-dehydrogenase (EC 1.1.1.49)	Dipeptide transport ATP-binding protein DppF (TC 3.A.1.5.2)	Dipeptide transport ATP-binding protein DppF (TC 3.A.1.5.2)	HPr kinase/phosphorylase (EC 2.7.1.-) (EC 2.7.4.-)
Glucose-6-phosphate isomerase (EC 5.3.1.9)	Dipeptide transport system permease protein DppB (TC 3.A.1.5.2)	Dipeptide transport system permease protein DppB (TC 3.A.1.5.2)	Hydrogen peroxide-inducible genes activator
Glucose-methanol-choline (GMC) oxidoreductase:NAD binding site	Diphosphomevalonate decarboxylase (EC 4.1.1.33)	Diphosphomevalonate decarboxylase (EC 4.1.1.33)	Hydroxyethylthiazole kinase (EC 2.7.1.50)
Glutamate 5-kinase (EC 2.7.2.11)	Dipicolinate synthase subunit B	Dipicolinate synthase subunit B	Hydroxymethylglutaryl-CoA lyase (EC 4.1.3.4)
Glutamate N-acetyltransferase (EC 2.3.1.35) / N-acetylglutamate synthase (EC 2.3.1.1)	DNA gyrase subunit A (EC 5.99.1.3)	DNA gyrase subunit A (EC 5.99.1.3)	Hydroxymethylglutaryl-CoA reductase (EC 1.1.1.34)
Glutamate synthase [NADPH] large chain (EC 1.4.1.13)	DNA gyrase subunit B (EC 5.99.1.3)	DNA gyrase subunit B (EC 5.99.1.3)	Hydroxymethylglutaryl-CoA synthase (EC 2.3.3.10)

Glutamate synthase [NADPH] small chain (EC 1.4.1.13)	DNA helicase (Rad25 homolog)	DNA helicase (Rad25 homolog)	Hydroxymethylpyrimidine phosphate synthase ThiC (EC 4.1.99.17)
Glutamate-1-semialdehyde aminotransferase (EC 5.4.3.8)	DNA ligase (EC 6.5.1.2)	DNA ligase (EC 6.5.1.2)	Hypothetical protein (associated with DNA helicase - Rad25 homolog)
Glutamine synthetase type I (EC 6.3.1.2)	DNA mismatch repair protein MutL	DNA mismatch repair protein MutL	Hypothetical transmembrane protein coupled to NADH-ubiquinone oxidoreductase chain 5 homolog
Glutaminyl-tRNA synthetase (EC 6.1.1.18)	DNA mismatch repair protein MutS	DNA mismatch repair protein MutS	Imidazoleglycerol-phosphate dehydratase (EC 4.2.1.19)
Glutamyl-tRNA(Gln) amidotransferase asparaginase subunit (EC 6.3.5.7)	DNA polymerase I (EC 2.7.7.7)	DNA polymerase I (EC 2.7.7.7)	Imidazolonepropionase (EC 3.5.2.7)
Glutamyl-tRNA(Gln) amidotransferase subunit A (EC 6.3.5.7)	DNA polymerase II (EC 2.7.7.7)	DNA polymerase II (EC 2.7.7.7)	IMP cyclohydrolase (EC 3.5.4.10)
Glutamyl-tRNA(Gln) amidotransferase subunit B (EC 6.3.5.7)	DNA polymerase III alpha subunit (EC 2.7.7.7)	DNA polymerase III alpha subunit (EC 2.7.7.7)	Indolepyruvate oxidoreductase subunit IorA (EC 1.2.7.8)
Glutamyl-tRNA(Gln) synthetase (EC 6.1.1.24)	DNA polymerase III beta subunit (EC 2.7.7.7)	DNA polymerase III beta subunit (EC 2.7.7.7)	Inorganic pyrophosphatase (EC 3.6.1.1)
Glutathione synthetase (EC 6.3.2.3)	DNA polymerase III delta prime subunit (EC 2.7.7.7)	DNA polymerase III delta prime subunit (EC 2.7.7.7)	Inosine-5'-monophosphate dehydrogenase (EC 1.1.1.205)
Glycerate kinase (EC 2.7.1.31)	DNA polymerase III delta subunit (EC 2.7.7.7)	DNA polymerase III delta subunit (EC 2.7.7.7)	Integration host factor beta subunit
Glycerol dehydrogenase (EC 1.1.1.6)	DNA polymerase III subunits gamma and tau (EC 2.7.7.7)	DNA polymerase III subunits gamma and tau (EC 2.7.7.7)	intein-containing
Glycerol kinase (EC 2.7.1.30)	DNA polymerase IV (EC 2.7.7.7)	DNA polymerase IV (EC 2.7.7.7)	Iron-sulfur cluster assembly protein SufB
Glycerol-3-phosphate ABC transporter, ATP-binding protein UgpC (TC 3.A.1.1.3)	DNA polymerase sliding clamp protein PCNA	DNA polymerase sliding clamp protein PCNA	Isocitrate dehydrogenase [NAD] (EC 1.1.1.41)
Glycerol-3-phosphate ABC transporter, periplasmic glycerol-3-phosphate-binding protein (TC 3.A.1.1.3)	DNA primase (EC 2.7.7.-)	DNA primase (EC 2.7.7.-)	Isocitrate dehydrogenase [NADP] (EC 1.1.1.42)
Glycerol-3-phosphate ABC transporter, permease protein UgpA (TC 3.A.1.1.3)	DNA primase small subunit (EC 2.7.7.-)	DNA primase small subunit (EC 2.7.7.-)	Isocitrate lyase (EC 4.1.3.1)
Glycerol-3-phosphate ABC transporter, permease protein UgpE (TC 3.A.1.1.3)	DNA recombination and repair protein RecF	DNA recombination and repair protein RecF	Isopentenyl-diphosphate delta-isomerase (EC 5.3.3.2)
Glycerol-3-phosphate dehydrogenase [NAD(P)+] (EC 1.1.1.94)	DNA recombination and repair protein RecO	DNA recombination and repair protein RecO	Kef-type K+ transport systems (NAD-binding component fused to domain related to exopolyphosphatase)
Glycine cleavage system H protein	DNA repair and recombination protein RadA	DNA repair and recombination protein RadA	Ketol-acid reductoisomerase (EC 1.1.1.86)
Glycine dehydrogenase [decarboxylating] (glycine cleavage system P protein) (EC 1.4.4.2)	DNA repair and recombination protein RadB	DNA repair and recombination protein RadB	Kynureninase (EC 3.7.1.3)
Glycine dehydrogenase [decarboxylating] (glycine cleavage system P2 protein) (EC 1.4.4.2)	DNA repair exonuclease family protein YhaO	DNA repair exonuclease family protein YhaO	L-lactate dehydrogenase (EC 1.1.1.27)

Glycogen phosphorylase (EC 2.4.1.1)	DNA repair protein RadC	DNA repair protein RadC	L-lactate permease
Glyoxylate reductase (EC 1.1.1.26)	DNA repair protein RecN	DNA repair protein RecN	L-proline glycine betaine ABC transport system permease protein ProV (TC 3.A.1.12.1)
GMP synthase [glutamine-hydrolyzing] (EC 6.3.5.2)	DNA topoisomerase I (EC 5.99.1.2)	DNA topoisomerase I (EC 5.99.1.2)	Lipoate synthase
GTP cyclohydrolase I (EC 3.5.4.16) type 1	DNA topoisomerase III (EC 5.99.1.2)	DNA topoisomerase III (EC 5.99.1.2)	Lipoate-protein ligase A
GTP cyclohydrolase II (EC 3.5.4.25)	DNA topoisomerase VI subunit A (EC 5.99.1.3)	DNA topoisomerase VI subunit A (EC 5.99.1.3)	Lipoate-protein ligase A, C-terminal 70 percent
GTP pyrophosphokinase (EC 2.7.6.5), (p)ppGpp synthetase II	DNA topoisomerase VI subunit B (EC 5.99.1.3)	DNA topoisomerase VI subunit B (EC 5.99.1.3)	Long-chain-fatty-acid--CoA ligase (EC 6.2.1.3)
GTP-binding and nucleic acid-binding protein YchF	DNA-3-methyladenine glycosylase II (EC 3.2.2.21)	DNA-3-methyladenine glycosylase II (EC 3.2.2.21)	Low-specificity L-threonine aldolase (EC 4.1.2.48)
GTP-binding protein EngB	DNA-binding protein HBsu	DNA-binding protein HBsu	LSU ribosomal protein L10Ae (L1p)
GTP-binding protein Era	DNA-binding protein HU	DNA-binding protein HU	LSU ribosomal protein L10e (L16p)
GTP-binding protein HflX	DNA-directed RNA polymerase alpha subunit (EC 2.7.7.6)	DNA-directed RNA polymerase alpha subunit (EC 2.7.7.6)	LSU ribosomal protein L11e (L5p)
GTPase and tRNA-U34 5-formylation enzyme TrmE	DNA-directed RNA polymerase beta' subunit (EC 2.7.7.6)	DNA-directed RNA polymerase beta' subunit (EC 2.7.7.6)	LSU ribosomal protein L11p (L12e)
Guanylate kinase (EC 2.7.4.8)	DNA-directed RNA polymerase delta subunit (EC 2.7.7.6)	DNA-directed RNA polymerase delta subunit (EC 2.7.7.6)	LSU ribosomal protein L12a (P1/P2)
Heat shock protein GrpE	DNA-directed RNA polymerase gamma subunit (EC 2.7.7.6)	DNA-directed RNA polymerase gamma subunit (EC 2.7.7.6)	LSU ribosomal protein L14p (L23e)
Heme A synthase, cytochrome oxidase biogenesis protein Cox15-CtaA	DNA-directed RNA polymerase subunit A' (EC 2.7.7.6)	DNA-directed RNA polymerase subunit A' (EC 2.7.7.6)	LSU ribosomal protein L15e
Heme O synthase, protoheme IX farnesyltransferase (EC 2.5.1.-) COX10-CtaB	DNA-directed RNA polymerase subunit A" (EC 2.7.7.6)	DNA-directed RNA polymerase subunit A" (EC 2.7.7.6)	LSU ribosomal protein L1e (L4p)
HflC protein	DNA-directed RNA polymerase subunit B' (EC 2.7.7.6)	DNA-directed RNA polymerase subunit B' (EC 2.7.7.6)	LSU ribosomal protein L20p
Histidine ammonia-lyase (EC 4.3.1.3)	DNA-directed RNA polymerase subunit B" (EC 2.7.7.6)	DNA-directed RNA polymerase subunit B" (EC 2.7.7.6)	LSU ribosomal protein L23Ae (L23p)
Histidinol dehydrogenase (EC 1.1.1.23)	DNA-directed RNA polymerase subunit D (EC 2.7.7.6)	DNA-directed RNA polymerase subunit D (EC 2.7.7.6)	LSU ribosomal protein L23e (L14p)
Histidinol-phosphate aminotransferase (EC 2.6.1.9)	DNA-directed RNA polymerase subunit E' (EC 2.7.7.6)	DNA-directed RNA polymerase subunit E' (EC 2.7.7.6)	LSU ribosomal protein L27Ae (L15p)
Holliday junction DNA helicase RuvB	DNA-directed RNA polymerase subunit F (EC 2.7.7.6)	DNA-directed RNA polymerase subunit F (EC 2.7.7.6)	LSU ribosomal protein L29p (L35e)
Homocitrate synthase (EC 2.3.3.14)	DNA-directed RNA polymerase subunit N (EC 2.7.7.6)	DNA-directed RNA polymerase subunit N (EC 2.7.7.6)	LSU ribosomal protein L2p (L8e)
Homoserine kinase (EC 2.7.1.39)	dTDP-4-dehydrorhamnose reductase (EC 1.1.1.133)	dTDP-4-dehydrorhamnose reductase (EC 1.1.1.133)	LSU ribosomal protein L31e
Homoserine O-acetyltransferase (EC 2.3.1.31)	dTDP-glucose 4,6-dehydratase (EC 4.2.1.46)	dTDP-glucose 4,6-dehydratase (EC 4.2.1.46)	LSU ribosomal protein L32e
HPr kinase/phosphorylase (EC 2.7.1.-) (EC 2.7.4.-)	Electron transfer flavoprotein, alpha subunit	Electron transfer flavoprotein, alpha subunit	LSU ribosomal protein L35e (L29p)
Hydrogen peroxide-inducible genes activator	Endonuclease III (EC 4.2.99.18)	Endonuclease III (EC 4.2.99.18)	LSU ribosomal protein L37Ae
Hydroxymethylglutaryl-CoA lyase (EC 4.1.3.4)	Endonuclease IV (EC 3.1.21.2)	Endonuclease IV (EC 3.1.21.2)	LSU ribosomal protein L3e (L3p)

Hydroxymethylglutaryl-CoA reductase (EC 1.1.1.34)	Enolase (EC 4.2.1.11)	Enolase (EC 4.2.1.11)	LSU ribosomal protein L44e
Hydroxymethylglutaryl-CoA synthase (EC 2.3.3.10)	Enoyl-[acyl-carrier-protein] reductase [FMN] (EC 1.3.1.9)	Enoyl-[acyl-carrier-protein] reductase [FMN] (EC 1.3.1.9)	LSU ribosomal protein L5e (L18p)
Hydroxymethylpyrimidine phosphate synthase ThiC (EC 4.1.99.17)	Enoyl-[acyl-carrier-protein] reductase [NADH] (EC 1.3.1.9)	Enoyl-[acyl-carrier-protein] reductase [NADH] (EC 1.3.1.9)	LSU ribosomal protein L7/L12 (P1/P2)
Hypothetical protein (associated with DNA helicase - Rad25 homolog)	Enoyl-CoA hydratase (EC 4.2.1.17)	Enoyl-CoA hydratase (EC 4.2.1.17)	LSU ribosomal protein L7Ae
Hypothetical transmembrane protein coupled to NADH-ubiquinone oxidoreductase chain 5 homolog	enoyl-CoA hydratase, R-specific	enoyl-CoA hydratase, R-specific	LSU ribosomal protein L8e (L2p)
Imidazole glycerol phosphate synthase amidotransferase subunit (EC 2.4.2.-)	Error-prone, lesion bypass DNA polymerase V (UmuC)	Error-prone, lesion bypass DNA polymerase V (UmuC)	LSU ribosomal protein P0 (L10p)
Imidazoleglycerol-phosphate dehydratase (EC 4.2.1.19)	Erythronate-4-phosphate dehydrogenase (EC 1.1.1.290)	Erythronate-4-phosphate dehydrogenase (EC 1.1.1.290)	Lysyl-tRNA synthetase (class I) (EC 6.1.1.6)
Imidazolonepropionase (EC 3.5.2.7)	Eukaryotic peptide chain release factor subunit 1	Eukaryotic peptide chain release factor subunit 1	Lysyl-tRNA synthetase (class II) (EC 6.1.1.6)
IMP cyclohydrolase (EC 3.5.4.10)	Eukaryotic translation initiation factor 2 alpha subunit	Eukaryotic translation initiation factor 2 alpha subunit	Macrolide export ATP-binding/permease protein MacB (EC 3.6.3.-)
Indole-3-glycerol phosphate synthase (EC 4.1.1.48)	Eukaryotic translation initiation factor 2 beta subunit	Eukaryotic translation initiation factor 2 beta subunit	Magnesium and cobalt efflux protein CorC
Indolepyruvate oxidoreductase subunit IorA (EC 1.2.7.8)	Eukaryotic translation initiation factor 2 gamma subunit	Eukaryotic translation initiation factor 2 gamma subunit	Malate dehydrogenase (EC 1.1.1.37)
Inorganic pyrophosphatase (EC 3.6.1.1)	Excinuclease ABC subunit A	Excinuclease ABC subunit A	Malate Na(+) symporter
Inosine-5'-monophosphate dehydrogenase (EC 1.1.1.205)	Excinuclease ABC subunit A paralog in greater Bacteroides group	Excinuclease ABC subunit A paralog in greater Bacteroides group	Maltodextrin glucosidase (EC 3.2.1.20)
Inosine-uridine preferring nucleoside hydrolase (EC 3.2.2.1)	Excinuclease ABC subunit B	Excinuclease ABC subunit B	Mannose-1-phosphate guanylyltransferase (GDP) (EC 2.7.7.22)
Integration host factor beta subunit	Excinuclease ABC subunit C	Excinuclease ABC subunit C	Mercuric ion reductase (EC 1.16.1.1)
intein-containing	Exodeoxyribonuclease III (EC 3.1.11.2)	Exodeoxyribonuclease III (EC 3.1.11.2)	Methenyltetrahydrofolate cyclohydrolase (EC 3.5.4.9)
Iron-sulfur cluster assembly protein SufB	Exodeoxyribonuclease VII large subunit (EC 3.1.11.6)	Exodeoxyribonuclease VII large subunit (EC 3.1.11.6)	Methylcrotonyl-CoA carboxylase biotin-containing subunit (EC 6.4.1.4)
Isocitrate dehydrogenase [NADP] (EC 1.1.1.42)	Exodeoxyribonuclease VII small subunit (EC 3.1.11.6)	Exodeoxyribonuclease VII small subunit (EC 3.1.11.6)	Methylcrotonyl-CoA carboxylase carboxyl transferase subunit (EC 6.4.1.4)
Isocitrate lyase (EC 4.1.3.1)	Exonuclease SbcC	Exonuclease SbcC	Methylenetetrahydrofolate dehydrogenase (NADP+) (EC 1.5.1.5)
Isopentenyl-diphosphate delta-isomerase (EC 5.3.3.2)	Exonuclease SbcD	Exonuclease SbcD	Methylmalonate-semialdehyde dehydrogenase (EC 1.2.1.27)
Kef-type K+ transport systems (NAD-binding component fused to domain related to exopolyphosphatase)	Expressed protein possibly involved in photorespiration	Expressed protein possibly involved in photorespiration	Methylmalonyl-CoA mutase (EC 5.4.99.2)
Ketoglutarate semialdehyde dehydrogenase (EC	Fe-S OXIDOREDUCTASE (1.8.-.-) Wyeosine biosynthesis	Fe-S OXIDOREDUCTASE (1.8.-.-) Wyeosine biosynthesis	Multi antimicrobial extrusion protein (Na(+)/drug antiporter), MATE

1.2.1.26)			family of MDR efflux pumps
Ketol-acid reductoisomerase (EC 1.1.1.86)	Ferredoxin	Ferredoxin	Multicopper oxidase
L-aspartate oxidase (EC 1.4.3.16)	Ferredoxin reductase	Ferredoxin reductase	N-acetyl-gamma-glutamyl-phosphate reductase (EC 1.2.1.38)
L-lactate dehydrogenase (EC 1.1.1.27)	Ferredoxin--NADP(+) reductase (EC 1.18.1.2)	Ferredoxin--NADP(+) reductase (EC 1.18.1.2)	N-acetyl-L,L-diaminopimelate deacetylase (EC 3.5.1.47)
L-lactate permease	Ferredoxin--nitrite reductase (EC 1.7.7.1)	Ferredoxin--nitrite reductase (EC 1.7.7.1)	N-acetylglutamate synthase (EC 2.3.1.1)
L-proline glycine betaine ABC transport system permease protein ProV (TC 3.A.1.12.1)	Ferric iron ABC transporter, ATP-binding protein	Ferric iron ABC transporter, ATP-binding protein	N-acyl homoserine lactone hydrolase
L-proline glycine betaine ABC transport system permease protein ProW (TC 3.A.1.12.1)	Ferric iron ABC transporter, permease protein	Ferric iron ABC transporter, permease protein	N-methylhydantoinase A (EC 3.5.2.14)
Light-independent protochlorophyllide reductase subunit B (EC 1.18.-.-)	Ferric siderophore transport system, biopolymer transport protein ExbB	Ferric siderophore transport system, biopolymer transport protein ExbB	Na(+) H(+) antiporter subunit D
Lipid A export ATP-binding/permease protein MsbA	Ferric siderophore transport system, periplasmic binding protein TonB	Ferric siderophore transport system, periplasmic binding protein TonB	NAD synthetase (EC 6.3.1.5)
Lipoate synthase	Ferrochelatase, protoheme ferro-lyase (EC 4.99.1.1)	Ferrochelatase, protoheme ferro-lyase (EC 4.99.1.1)	NAD-dependent glyceraldehyde-3-phosphate dehydrogenase (EC 1.2.1.12)
Lipoprotein signal peptidase (EC 3.4.23.36)	Fibrillarin	Fibrillarin	NAD-specific glutamate dehydrogenase (EC 1.4.1.2)
Long-chain-fatty-acid--CoA ligase (EC 6.2.1.3)	FIG000605: protein co-occurring with transport systems (COG1739)	FIG000605: protein co-occurring with transport systems (COG1739)	NAD(P)-dependent glyceraldehyde 3-phosphate dehydrogenase archaeal (EC 1.2.1.59)
low-specificity D-threonine aldolase	Flagella-related protein FlaE	Flagella-related protein FlaE	NADH dehydrogenase (EC 1.6.99.3)
Low-specificity L-threonine aldolase (EC 4.1.2.48)	Flagella-related protein FlaH	Flagella-related protein FlaH	NADH-ubiquinone oxidoreductase chain B (EC 1.6.5.3)
LSU ribosomal protein L10Ae (L1p)	Flagella-related protein FlaI	Flagella-related protein FlaI	NADH-ubiquinone oxidoreductase chain D (EC 1.6.5.3)
LSU ribosomal protein L10e (L16p)	Flagella-related protein FlaJ	Flagella-related protein FlaJ	NADH-ubiquinone oxidoreductase chain E (EC 1.6.5.3)
LSU ribosomal protein L11p (L12e)	Flagellar assembly protein FliH	Flagellar assembly protein FliH	NADH-ubiquinone oxidoreductase chain G (EC 1.6.5.3)
LSU ribosomal protein L12e (L11p)	Flagellar basal-body rod protein FlgC	Flagellar basal-body rod protein FlgC	NADH-ubiquinone oxidoreductase chain H (EC 1.6.5.3)
LSU ribosomal protein L16p (L10e)	Flagellar basal-body rod protein FlgF	Flagellar basal-body rod protein FlgF	NADH-ubiquinone oxidoreductase chain I (EC 1.6.5.3)
LSU ribosomal protein L19p	Flagellar biosynthesis protein FlhA	Flagellar biosynthesis protein FlhA	NADH-ubiquinone oxidoreductase chain L (EC 1.6.5.3)
LSU ribosomal protein L1p (L10Ae)	Flagellar biosynthesis protein FlhA	Flagellar biosynthesis protein FlhA	NADH-ubiquinone oxidoreductase chain M (EC 1.6.5.3)

LSU ribosomal protein L20p	Flagellar biosynthesis protein FlhB	Flagellar biosynthesis protein FlhB	NADH-ubiquinone oxidoreductase chain N (EC 1.6.5.3)
LSU ribosomal protein L21e	Flagellar biosynthesis protein FlhF	Flagellar biosynthesis protein FlhF	NADP-dependent malic enzyme (EC 1.1.1.40)
LSU ribosomal protein L24p (L26e)	Flagellar biosynthesis protein FliP	Flagellar biosynthesis protein FliP	Naphthoate synthase (EC 4.1.3.36)
LSU ribosomal protein L27Ae (L15p)	Flagellar biosynthesis protein FliR	Flagellar biosynthesis protein FliR	Neopullulanase (EC 3.2.1.135)
LSU ribosomal protein L2p (L8e)	Flagellar hook protein FlgE	Flagellar hook protein FlgE	Nickel responsive regulator NikR
LSU ribosomal protein L37e	Flagellar hook-associated protein FlgK	Flagellar hook-associated protein FlgK	Nicotinamidase (EC 3.5.1.19)
LSU ribosomal protein L5p (L11e)	Flagellar hook-associated protein FliD	Flagellar hook-associated protein FliD	Nicotinate phosphoribosyltransferase (EC 2.4.2.11)
LSU ribosomal protein L7Ae	Flagellar M-ring protein FliF	Flagellar M-ring protein FliF	Nitric-oxide reductase (EC 1.7.99.7), quinol-dependent
LSU ribosomal protein L8e (L2p)	Flagellar motor switch protein FliM	Flagellar motor switch protein FliM	Nitrite reductase probable [NAD(P)H] subunit (EC 1.7.1.4)
LSU ribosomal protein L9e (L6p)	Flagellar motor switch protein FliN	Flagellar motor switch protein FliN	Nucleoside diphosphate kinase (EC 2.7.4.6)
LSU ribosomal protein L9p	Flagellin protein FlaA	Flagellin protein FlaA	Nucleoside permease NupC
LSU ribosomal protein L9p	Flagellum-specific ATP synthase FliI	Flagellum-specific ATP synthase FliI	O-succinylbenzoic acid--CoA ligase (EC 6.2.1.26)
LSU ribosomal protein P0 (L10p)	Flap structure-specific endonuclease (EC 3.-.-.-)	Flap structure-specific endonuclease (EC 3.-.-.-)	Origin of replication recognition protein
Lysine 2,3-aminomutase (EC 5.4.3.2)	FMN adenylyltransferase, type 3 archaeal (EC 2.7.7.2)	FMN adenylyltransferase, type 3 archaeal (EC 2.7.7.2)	Ornithine carbamoyltransferase (EC 2.1.3.3)
Macrolide export ATP-binding/permease protein MacB (EC 3.6.3.-)	Folate-dependent protein for Fe/S cluster synthesis/repair in oxidative stress	Folate-dependent protein for Fe/S cluster synthesis/repair in oxidative stress	Ornithine cyclodeaminase (EC 4.3.1.12)
Magnesium and cobalt efflux protein CorC	Folylpolyglutamate synthase (EC 6.3.2.17)	Folylpolyglutamate synthase (EC 6.3.2.17)	Orotate phosphoribosyltransferase (EC 2.4.2.10)
Malate dehydrogenase (EC 1.1.1.37)	Formamidopyrimidine-DNA glycosylase (EC 3.2.2.23)	Formamidopyrimidine-DNA glycosylase (EC 3.2.2.23)	oxidoreductase of aldo/keto reductase family, subgroup 1
Malate Na(+) symporter	Formate dehydrogenase-O, iron-sulfur subunit (EC 1.2.1.2)	Formate dehydrogenase-O, iron-sulfur subunit (EC 1.2.1.2)	Pantoate kinase, archaeal (EC 2.7.1.-)
Malto-oligosyltrehalose trehalohydrolase (EC 3.2.1.141)	Formate dehydrogenase-O, major subunit (EC 1.2.1.2)	Formate dehydrogenase-O, major subunit (EC 1.2.1.2)	Pantothenate:Na+ symporter (TC 2.A.21.1.1)
Maltodextrin glucosidase (EC 3.2.1.20)	Formate--tetrahydrofolate ligase (EC 6.3.4.3)	Formate--tetrahydrofolate ligase (EC 6.3.4.3)	Para-aminobenzoate synthase, amidotransferase component (EC 2.6.1.85)
Maltose O-acetyltransferase (EC 2.3.1.79)	Formiminoglutamic iminohydrolase (EC 3.5.3.13)	Formiminoglutamic iminohydrolase (EC 3.5.3.13)	Peptidase M48, Ste24p precursor
Maltose/maltodextrin transport ATP-binding protein MalK (EC 3.6.3.19)	Fructose-1,6-bisphosphatase, GlpX type (EC 3.1.3.11)	Fructose-1,6-bisphosphatase, GlpX type (EC 3.1.3.11)	Peptide chain release factor 1
Manganese superoxide dismutase (EC 1.15.1.1)	Fructose-1,6-bisphosphatase, type I (EC 3.1.3.11)	Fructose-1,6-bisphosphatase, type I (EC 3.1.3.11)	Peptide deformylase (EC 3.5.1.88)
Mannose-1-phosphate guanylyltransferase (EC 2.7.7.13)	Fructose-bisphosphate aldolase class I (EC 4.1.2.13)	Fructose-bisphosphate aldolase class I (EC 4.1.2.13)	Peptide methionine sulfoxide reductase MsrA (EC 1.8.4.11)
Membrane-associated zinc metalloprotease	Fructose-bisphosphate aldolase, archaeal class I (EC	Fructose-bisphosphate aldolase, archaeal class I (EC 4.1.2.13)	Peptidyl-prolyl cis-trans isomerase (EC 5.2.1.8)

	4.1.2.13)		
Mercuric ion reductase (EC 1.16.1.1)	Fumarate hydratase class II (EC 4.2.1.2)	Fumarate hydratase class II (EC 4.2.1.2)	Permease of the drug/metabolite transporter (DMT) superfamily
Methenyltetrahydrofolate cyclohydrolase (EC 3.5.4.9)	Fumarylacetoacetase (EC 3.7.1.2)	Fumarylacetoacetase (EC 3.7.1.2)	PF00070 family, FAD-dependent NAD(P)-disulphide oxidoreductase
Methionyl-tRNA formyltransferase (EC 2.1.2.9)	Fumarylacetoacetate hydrolase family protein	Fumarylacetoacetate hydrolase family protein	Phage tail fiber protein
Methylcrotonyl-CoA carboxylase carboxyl transferase subunit (EC 6.4.1.4)	**G:T/U mismatch-specific uracil/thymine DNA-glycosylase**	G:T/U mismatch-specific uracil/thymine DNA-glycosylase	Phenylalanyl-tRNA synthetase alpha chain (EC 6.1.1.20)
Methylenetetrahydrofolate dehydrogenase (NADP+) (EC 1.5.1.5)	Gamma-glutamyl phosphate reductase (EC 1.2.1.41)	Gamma-glutamyl phosphate reductase (EC 1.2.1.41)	Phenylalanyl-tRNA synthetase beta chain (EC 6.1.1.20)
Methylglutaconyl-CoA hydratase (EC 4.2.1.18)	GDP-mannose mannosyl hydrolase (EC 3.6.1.-)	GDP-mannose mannosyl hydrolase (EC 3.6.1.-)	Phosphate regulon transcriptional regulatory protein PhoB (SphR)
Methylglyoxal synthase (EC 4.2.3.3)	Geranyl-CoA carboxylase biotin-containing subunit (EC 6.4.1.5)	Geranyl-CoA carboxylase biotin-containing subunit (EC 6.4.1.5)	Phosphate starvation-inducible protein PhoH, predicted ATPase
Methylisocitrate lyase (EC 4.1.3.30)	Germination-specific N-acetylmuramoyl-L-alanine amidase (EC 3.5.1.28), cell wall hydrolase CwlD	Germination-specific N-acetylmuramoyl-L-alanine amidase (EC 3.5.1.28), cell wall hydrolase CwlD	Phosphate transport ATP-binding protein PstB (TC 3.A.1.7.1)
Methylmalonyl-CoA mutase (EC 5.4.99.2)	gliding motility protein GldF	gliding motility protein GldF	Phosphate transport system permease protein PstA (TC 3.A.1.7.1)
Methylthioadenosine deaminase	Glucoamylase (EC 3.2.1.3)	Glucoamylase (EC 3.2.1.3)	Phosphate transport system regulatory protein PhoU
Multiple sugar ABC transporter, ATP-binding protein	Gluconate dehydratase (EC 4.2.1.39)	Gluconate dehydratase (EC 4.2.1.39)	Phosphoadenylyl-sulfate reductase [thioredoxin] (EC 1.8.4.8)
N-acetyl-gamma-glutamyl-phosphate reductase (EC 1.2.1.38)	Glucosamine--fructose-6-phosphate aminotransferase [isomerizing] (EC 2.6.1.16)	Glucosamine--fructose-6-phosphate aminotransferase [isomerizing] (EC 2.6.1.16)	Phosphoenolpyruvate carboxykinase [ATP] (EC 4.1.1.49)
N-acetyl-L,L-diaminopimelate deacetylase (EC 3.5.1.47)	Glucosamine-6-phosphate deaminase (EC 3.5.99.6)	Glucosamine-6-phosphate deaminase (EC 3.5.99.6)	Phosphoenolpyruvate carboxylase (EC 4.1.1.31)
N-acetylglucosamine-1-phosphate uridyltransferase (EC 2.7.7.23)	Glucose-1-phosphate adenylyltransferase (EC 2.7.7.27)	Glucose-1-phosphate adenylyltransferase (EC 2.7.7.27)	Phosphoenolpyruvate synthase (EC 2.7.9.2)
N-acyl homoserine lactone hydrolase	Glucose-1-phosphate cytidylyltransferase (EC 2.7.7.33)	Glucose-1-phosphate cytidylyltransferase (EC 2.7.7.33)	Phosphoglycerate kinase (EC 2.7.2.3)
N-methylhydantoinase (ATP-hydrolyzing) (EC 3.5.2.14)	Glucose-1-phosphate thymidylyltransferase (EC 2.7.7.24)	Glucose-1-phosphate thymidylyltransferase (EC 2.7.7.24)	Phosphopantetheine adenylyltransferase, type II eukaryotic (EC 2.7.7.3)
N-methylhydantoinase A (EC 3.5.2.14)	Glucose-6-phosphate 1-dehydrogenase (EC 1.1.1.49)	Glucose-6-phosphate 1-dehydrogenase (EC 1.1.1.49)	Phosphopantothenate synthetase, archaeal
N-methylhydantoinase B (EC 3.5.2.14)	Glucose-6-phosphate isomerase (EC 5.3.1.9)	Glucose-6-phosphate isomerase (EC 5.3.1.9)	Phosphoribosyl-AMP cyclohydrolase (EC 3.5.4.19)
Na(+) H(+) antiporter subunit D	Glucose-methanol-choline (GMC) oxidoreductase:NAD binding site	Glucose-methanol-choline (GMC) oxidoreductase:NAD binding site	Phosphoribosylaminoimidazole carboxylase ATPase subunit (EC 4.1.1.21)
Na+/H+ antiporter NhaA type	Glutamate 5-kinase (EC 2.7.2.11)	Glutamate 5-kinase (EC 2.7.2.11)	Phosphoribosylaminoimidazole-succinocarboxamide synthase (EC 6.3.2.6)
NAD kinase (EC 2.7.1.23)	Glutamate racemase (EC 5.1.1.3)	Glutamate racemase (EC 5.1.1.3)	Phosphoribosylformimino-5-aminoimidazole carboxamide ribotide isomerase (EC 5.3.1.16)

NAD synthetase (EC 6.3.1.5)
NAD-dependent formate dehydrogenase alpha subunit
NAD-dependent glyceraldehyde-3-phosphate dehydrogenase (EC 1.2.1.12)
NAD-independent protein deacetylase AcuC
NAD-reducing hydrogenase subunit HoxF (EC 1.12.1.2)
NAD-specific glutamate dehydrogenase (EC 1.4.1.2)
NAD(P) transhydrogenase alpha subunit (EC 1.6.1.2)
NAD(P) transhydrogenase subunit beta (EC 1.6.1.2)
NAD(P)-dependent glyceraldehyde 3-phosphate dehydrogenase archaeal (EC 1.2.1.59)
NADH dehydrogenase (EC 1.6.99.3)
NADH dehydrogenase, subunit 5
NADH ubiquinone oxidoreductase chain A (EC 1.6.5.3)
NADH-ubiquinone oxidoreductase chain B (EC 1.6.5.3)
NADH-ubiquinone oxidoreductase chain C (EC 1.6.5.3)
NADH-ubiquinone oxidoreductase chain D (EC 1.6.5.3)
NADH-ubiquinone oxidoreductase chain E (EC 1.6.5.3)
NADH-ubiquinone oxidoreductase chain G (EC 1.6.5.3)
NADH-ubiquinone oxidoreductase chain H (EC 1.6.5.3)
NADH-ubiquinone oxidoreductase chain I (EC 1.6.5.3)
NADH-ubiquinone oxidoreductase chain K (EC

Glutamate synthase [NADPH] large chain (EC 1.4.1.13)
Glutamate--cysteine ligase archaeal (EC 6.3.2.2)
Glutamate-1-semialdehyde aminotransferase (EC 5.4.3.8)
Glutaminase (EC 3.5.1.2)
Glutamine synthetase type I (EC 6.3.1.2)
Glutamine-dependent 2-keto-4-methylthiobutyrate transaminase
Glutaminyl-tRNA synthetase (EC 6.1.1.18)
Glutamyl-tRNA(Gln) amidotransferase asparaginase subunit (EC 6.3.5.7)
Glutamyl-tRNA(Gln) amidotransferase subunit A (EC 6.3.5.7)
Glutamyl-tRNA(Gln) amidotransferase subunit B (EC 6.3.5.7)
Glutamyl-tRNA(Gln) amidotransferase transferase subunit (EC 6.3.5.7)
Glutamyl-tRNA(Gln) synthetase (EC 6.1.1.24)
Glycerol dehydrogenase (EC 1.1.1.6)
Glycerol kinase (EC 2.7.1.30)
Glycerol-1-phosphate dehydrogenase [NAD(P)] (EC 1.1.1.261)
Glycerol-3-phosphate ABC transporter, periplasmic glycerol-3-phosphate-binding protein (TC 3.A.1.1.3)
Glycerol-3-phosphate ABC transporter, permease protein UgpA (TC 3.A.1.1.3)
Glycerol-3-phosphate ABC transporter, permease protein UgpE (TC 3.A.1.1.3)
Glycerol-3-phosphate dehydrogenase (EC 1.1.5.3)
Glycerol-3-phosphate dehydrogenase [NAD(P)+] (EC

Glutamate synthase [NADPH] large chain (EC 1.4.1.13)
Glutamate--cysteine ligase archaeal (EC 6.3.2.2)
Glutamate-1-semialdehyde aminotransferase (EC 5.4.3.8)
Glutaminase (EC 3.5.1.2)
Glutamine synthetase type I (EC 6.3.1.2)
Glutamine-dependent 2-keto-4-methylthiobutyrate transaminase
Glutaminyl-tRNA synthetase (EC 6.1.1.18)
Glutamyl-tRNA(Gln) amidotransferase asparaginase subunit (EC 6.3.5.7)
Glutamyl-tRNA(Gln) amidotransferase subunit A (EC 6.3.5.7)
Glutamyl-tRNA(Gln) amidotransferase subunit B (EC 6.3.5.7)
Glutamyl-tRNA(Gln) amidotransferase transferase subunit (EC 6.3.5.7)
Glutamyl-tRNA(Gln) synthetase (EC 6.1.1.24)
Glycerol dehydrogenase (EC 1.1.1.6)
Glycerol kinase (EC 2.7.1.30)
Glycerol-1-phosphate dehydrogenase [NAD(P)] (EC 1.1.1.261)
Glycerol-3-phosphate ABC transporter, periplasmic glycerol-3-phosphate-binding protein (TC 3.A.1.1.3)
Glycerol-3-phosphate ABC transporter, permease protein UgpA (TC 3.A.1.1.3)
Glycerol-3-phosphate ABC transporter, permease protein UgpE (TC 3.A.1.1.3)
Glycerol-3-phosphate dehydrogenase (EC 1.1.5.3)
Glycerol-3-phosphate dehydrogenase [NAD(P)+] (EC 1.1.1.94)

Phosphoribosylformylglycinamidine cyclo-ligase (EC 6.3.3.1)
Phosphoribosylformylglycinamidine synthase, glutamine amidotransferase subunit (EC 6.3.5.3)
Phosphoribosylformylglycinamidine synthase, synthetase subunit (EC 6.3.5.3)
Phosphoribosylglycinamide formyltransferase (EC 2.1.2.2)
Phosphoserine phosphatase (EC 3.1.3.3)
Phytoene dehydrogenase (EC 1.14.99.-)
Phytoene desaturase (EC 1.14.99.-)
Phytoene synthase (EC 2.5.1.32)
Polyphosphate kinase (EC 2.7.4.1)
Porphobilinogen deaminase (EC 2.5.1.61)
Potassium channel protein
Potassium efflux system KefA protein
Potassium uptake protein TrkH
Potassium uptake protein, integral membrane component, KtrB
Predicted L-lactate dehydrogenase, Fe-S oxidoreductase subunit YkgE
Predicted N6-adenine-specific RNA methylase containing THUMP domain
Prephenate dehydratase (EC 4.2.1.51)
Probable electron transfer flavoprotein-quinone oxidoreductase FixC (EC 1.5.5.-)
Probable low-affinity inorganic phosphate transporter
Probable peroxiredoxin (EC 1.11.1.15)

1.6.5.3)
NADH-ubiquinone oxidoreductase chain L (EC 1.6.5.3)
NADH-ubiquinone oxidoreductase chain L (EC 1.6.5.3)
NADH-ubiquinone oxidoreductase chain M (EC 1.6.5.3)
NADH-ubiquinone oxidoreductase chain N (EC 1.6.5.3)
NADP-dependent malic enzyme (EC 1.1.1.40)
NADP-specific glutamate dehydrogenase (EC 1.4.1.4)
NADPH-dependent glyceraldehyde-3-phosphate dehydrogenase (EC 1.2.1.13)
Naphthoate synthase (EC 4.1.3.36)
Neopullulanase (EC 3.2.1.135)
Nickel transporter UreH
Nicotinamidase (EC 3.5.1.19)
Nicotinate phosphoribosyltransferase (EC 2.4.2.11)
Nicotinate-nucleotide adenylyltransferase (EC 2.7.7.18)
Nitrate/nitrite transporter
Nitric oxide -responding transcriptional regulator NnrR (Crp/Fnr family)
Nitric-oxide reductase (EC 1.7.99.7), quinol-dependent
Nitrite reductase probable [NAD(P)H] subunit (EC 1.7.1.4)
Nitrogen regulatory protein P-II
Nitrous oxide reductase maturation protein NosF (ATPase)

1.1.1.94)
Glycerophosphoryl diester phosphodiesterase (EC 3.1.4.46)
Glycine betaine ABC transport system, glycine betaine-binding protein OpuAC
Glycine betaine ABC transport system, permease protein OpuAB
Glycine betaine transporter OpuD
Glycine cleavage system H protein
Glycine dehydrogenase [decarboxylating] (glycine cleavage system P protein) (EC 1.4.4.2)
Glycine dehydrogenase [decarboxylating] (glycine cleavage system P1 protein) (EC 1.4.4.2)
Glycine dehydrogenase [decarboxylating] (glycine cleavage system P2 protein) (EC 1.4.4.2)
Glycine N-methyltransferase (EC 2.1.1.20)
Glycine oxidase ThiO (EC 1.4.3.19)
Glycogen phosphorylase (EC 2.4.1.1)
Glycogen synthase, ADP-glucose transglucosylase (EC 2.4.1.21)
Glycolate dehydrogenase (EC 1.1.99.14), iron-sulfur subunit GlcF
Glycolate dehydrogenase (EC 1.1.99.14), subunit GlcD
Glyoxylate reductase (EC 1.1.1.26)
GMP reductase (EC 1.7.1.7)
GMP synthase [glutamine-hydrolyzing] (EC 6.3.5.2)
GTP cyclohydrolase I (EC 3.5.4.16) type 1
GTP cyclohydrolase I (EC 3.5.4.16) type 2

Glycerophosphoryl diester phosphodiesterase (EC 3.1.4.46)
Glycine betaine ABC transport system, glycine betaine-binding protein OpuAC
Glycine betaine ABC transport system, permease protein OpuAB
Glycine betaine transporter OpuD
Glycine cleavage system H protein
Glycine dehydrogenase [decarboxylating] (glycine cleavage system P protein) (EC 1.4.4.2)
Glycine dehydrogenase [decarboxylating] (glycine cleavage system P1 protein) (EC 1.4.4.2)
Glycine dehydrogenase [decarboxylating] (glycine cleavage system P2 protein) (EC 1.4.4.2)
Glycine N-methyltransferase (EC 2.1.1.20)
Glycine oxidase ThiO (EC 1.4.3.19)
Glycogen phosphorylase (EC 2.4.1.1)
Glycogen synthase, ADP-glucose transglucosylase (EC 2.4.1.21)
Glycolate dehydrogenase (EC 1.1.99.14), iron-sulfur subunit GlcF
Glycolate dehydrogenase (EC 1.1.99.14), subunit GlcD
Glyoxylate reductase (EC 1.1.1.26)
GMP reductase (EC 1.7.1.7)
GMP synthase [glutamine-hydrolyzing] (EC 6.3.5.2)
GTP cyclohydrolase I (EC 3.5.4.16) type 1
GTP cyclohydrolase I (EC 3.5.4.16) type 2

Propionyl-CoA carboxylase beta chain (EC 6.4.1.3)
Proteasome subunit alpha (EC 3.4.25.1), archaeal
Proteasome subunit beta (EC 3.4.25.1), archaeal
Proteasome-activating AAA-ATPase (PAN), archaeal
Putative formate dehydrogenase oxidoreductase protein
Putative periplasmic phosphate-binding protein PstS (Halobacteriales type)
Putative stomatin/prohibitin-family membrane protease subunit aq_911
Putative Zn-dependent hydrolase in polyisoprenoid biosynthetic cluster
Putrescine transport ATP-binding protein PotA (TC 3.A.1.11.1)
Pyridoxine biosynthesis glutamine amidotransferase, glutaminase subunit (EC 2.4.2.-)
Pyridoxine biosynthesis glutamine amidotransferase, synthase subunit (EC 2.4.2.-)
Pyrimidine-nucleoside phosphorylase (EC 2.4.2.2)
Pyruvate dehydrogenase E1 component alpha subunit (EC 1.2.4.1)
Pyruvate kinase (EC 2.7.1.40)
Pyruvate-flavodoxin oxidoreductase (EC 1.2.7.-)
Pyruvate,phosphate dikinase (EC 2.7.9.1)
Pyruvate:ferredoxin oxidoreductase, beta subunit (EC 1.2.7.1)
Pyruvate:ferredoxin oxidoreductase, delta subunit (EC 1.2.7.1)
Quinolinate synthetase (EC 2.5.1.72)

Nitrous-oxide reductase (EC 1.7.99.6)	GTP cyclohydrolase II (EC 3.5.4.25)	GTP cyclohydrolase II (EC 3.5.4.25)	Radical SAM domain heme biosynthesis protein
Nucleoside diphosphate kinase (EC 2.7.4.6)	GTP pyrophosphokinase (EC 2.7.6.5), (p)ppGpp synthetase II	GTP pyrophosphokinase (EC 2.7.6.5), (p)ppGpp synthetase II	RecA protein
Nucleoside permease NupC	GTP-binding and nucleic acid-binding protein YchF	GTP-binding and nucleic acid-binding protein YchF	Replication factor C large subunit
Nucleoside triphosphate pyrophosphohydrolase MazG (EC 3.6.1.8)	GTP-binding protein EngB	GTP-binding protein EngB	Respiratory nitrate reductase alpha chain (EC 1.7.99.4)
Oligopeptide ABC transporter, periplasmic oligopeptide-binding protein OppA (TC 3.A.1.5.1)	GTP-binding protein Era	GTP-binding protein Era	Respiratory nitrate reductase beta chain (EC 1.7.99.4)
Oligopeptide transport system permease protein OppB (TC 3.A.1.5.1)	GTP-binding protein HflX	GTP-binding protein HflX	Rhamnulokinase (EC 2.7.1.5)
Origin of replication recognition protein	GTPase and tRNA-U34 5-formylation enzyme TrmE	GTPase and tRNA-U34 5-formylation enzyme TrmE	Rhodanese domain protein, Enterobacterial subgroup, YceA homolog
Ornithine carbamoyltransferase (EC 2.1.3.3)	Guanine-hypoxanthine permease	Guanine-hypoxanthine permease	Rhodanese-like domain required for thiamine synthesis
Ornithine cyclodeaminase (EC 4.3.1.12)	Guanylate kinase (EC 2.7.4.8)	Guanylate kinase (EC 2.7.4.8)	Riboflavin synthase eubacterial/eukaryotic (EC 2.5.1.9)
Orotidine 5'-phosphate decarboxylase (EC 4.1.1.23)	Heat shock protein GrpE	Heat shock protein GrpE	Ribokinase (EC 2.7.1.15)
Pantoate--beta-alanine ligase (EC 6.3.2.1)	Heat-inducible transcription repressor HrcA	Heat-inducible transcription repressor HrcA	Ribonuclease HII (EC 3.1.26.4)
Pantothenate:Na+ symporter (TC 2.A.21.1.1)	Helicase loader DnaI	Helicase loader DnaI	Ribonuclease P protein component (EC 3.1.26.5)
Para-aminobenzoate synthase, amidotransferase component (EC 2.6.1.85)	Heme A synthase, cytochrome oxidase biogenesis protein Cox15-CtaA	Heme A synthase, cytochrome oxidase biogenesis protein Cox15-CtaA	Ribonuclease P protein component 2 (EC 3.1.26.5)
Penicillin-binding protein 2 (PBP-2)	Heme biosynthesis protein related to NirG	Heme biosynthesis protein related to NirG	Ribonuclease Z (EC 3.1.26.11)
Peptidase M48, Ste24p precursor	Heme O synthase, protoheme IX farnesyltransferase (EC 2.5.1.-) COX10-CtaB	Heme O synthase, protoheme IX farnesyltransferase (EC 2.5.1.-) COX10-CtaB	Ribonucleotide reductase of class Ia (aerobic), alpha subunit (EC 1.17.4.1)
Peptide deformylase (EC 3.5.1.88)	Heptaprenyl diphosphate synthase component II (EC 2.5.1.30)	Heptaprenyl diphosphate synthase component II (EC 2.5.1.30)	Ribonucleotide reductase of class Ia (aerobic), beta subunit (EC 1.17.4.1)
Peptide methionine sulfoxide reductase MsrA (EC 1.8.4.11)	Hexokinase (EC 2.7.1.1)	Hexokinase (EC 2.7.1.1)	Ribonucleotide reductase of class II (coenzyme B12-dependent) (EC 1.17.4.1)
Peptidyl-prolyl cis-trans isomerase ppiA precursor (EC 5.2.1.8)	High-affinity branched-chain amino acid transport system permease protein LivH (TC 3.A.1.4.1)	High-affinity branched-chain amino acid transport system permease protein LivH (TC 3.A.1.4.1)	Ribose-1,5-bisphosphate isomerase
Peptidyl-prolyl cis-trans isomerase ppiD (EC 5.2.1.8)	Histidinol dehydrogenase (EC 1.1.1.23)	Histidinol dehydrogenase (EC 1.1.1.23)	Ribose-phosphate pyrophosphokinase (EC 2.7.6.1)
PF00070 family, FAD-dependent NAD(P)-disulphide oxidoreductase	Histidinol-phosphatase (EC 3.1.3.15)	Histidinol-phosphatase (EC 3.1.3.15)	Ribosomal large subunit pseudouridine synthase B (EC 4.2.1.70)
Phage tail fiber protein	Histidinol-phosphate aminotransferase (EC 2.6.1.9)	Histidinol-phosphate aminotransferase (EC 2.6.1.9)	RNA polymerase sigma-54 factor RpoN
Phenylalanine-4-hydroxylase (EC 1.14.16.1)	Holliday junction DNA helicase RuvA	Holliday junction DNA helicase RuvA	Rod shape-determining protein RodA
Phenylalanyl-tRNA synthetase alpha chain (EC	Holliday junction DNA helicase RuvB	Holliday junction DNA helicase RuvB	rRNA biogenesis protein Nop5/Nop56

6.1.1.20)			
Phenylalanyl-tRNA synthetase beta chain (EC 6.1.1.20)	Holo-[acyl-carrier protein] synthase (EC 2.7.8.7)	Holo-[acyl-carrier protein] synthase (EC 2.7.8.7)	rRNA small subunit methyltransferase I
PhnJ protein	Homogentisate 1,2-dioxygenase (EC 1.13.11.5)	Homogentisate 1,2-dioxygenase (EC 1.13.11.5)	S-adenosylhomocysteine deaminase (EC 3.5.4.28)
Phosphate ABC transporter, periplasmic phosphate-binding protein PstS (TC 3.A.1.7.1)	Homoserine dehydrogenase (EC 1.1.1.3)	Homoserine dehydrogenase (EC 1.1.1.3)	S-adenosylmethionine synthetase (EC 2.5.1.6)
Phosphate regulon transcriptional regulatory protein PhoB (SphR)	Homoserine kinase (EC 2.7.1.39)	Homoserine kinase (EC 2.7.1.39)	S-adenosylmethionine:tRNA ribosyltransferase-isomerase (EC 5.-.-.-)
Phosphate transport ATP-binding protein PstB (TC 3.A.1.7.1)	Homoserine O-acetyltransferase (EC 2.3.1.31)	Homoserine O-acetyltransferase (EC 2.3.1.31)	Serine acetyltransferase (EC 2.3.1.30)
Phosphate transport system permease protein PstA (TC 3.A.1.7.1)	HPr kinase/phosphorylase (EC 2.7.1.-) (EC 2.7.4.-)	HPr kinase/phosphorylase (EC 2.7.1.-) (EC 2.7.4.-)	Serine hydroxymethyltransferase (EC 2.1.2.1)
Phosphate transport system permease protein PstC (TC 3.A.1.7.1)	Hydrogen peroxide-inducible genes activator	Hydrogen peroxide-inducible genes activator	Serine protease precursor MucD/AlgY associated with sigma factor RpoE
Phosphate transport system regulatory protein PhoU	Hydroxyacylglutathione hydrolase (EC 3.1.2.6)	Hydroxyacylglutathione hydrolase (EC 3.1.2.6)	Serine-pyruvate aminotransferase/archaeal aspartate aminotransferase
Phosphatidate cytidylyltransferase (EC 2.7.7.41)	Hydroxymethylglutaryl-CoA reductase (EC 1.1.1.34)	Hydroxymethylglutaryl-CoA reductase (EC 1.1.1.34)	Seryl-tRNA synthetase (EC 6.1.1.11)
Phosphoadenylyl-sulfate reductase [thioredoxin] (EC 1.8.4.8)	Hydroxymethylglutaryl-CoA synthase (EC 2.3.3.10)	Hydroxymethylglutaryl-CoA synthase (EC 2.3.3.10)	Signal recognition particle, subunit Ffh SRP54 (TC 3.A.5.1.1)
Phosphoenolpyruvate carboxykinase [ATP] (EC 4.1.1.49)	Hydroxymethylpyrimidine ABC transporter, ATPase component	Hydroxymethylpyrimidine ABC transporter, ATPase component	SOS-response repressor and protease LexA (EC 3.4.21.88)
Phosphoenolpyruvate carboxylase (EC 4.1.1.31)	Hydroxymethylpyrimidine ABC transporter, substrate-binding component	Hydroxymethylpyrimidine ABC transporter, substrate-binding component	Spermidine Putrescine ABC transporter permease component potC (TC_3.A.1.11.1)
Phosphoenolpyruvate synthase (EC 2.7.9.2)	Hydroxymethylpyrimidine ABC transporter, transmembrane component	Hydroxymethylpyrimidine ABC transporter, transmembrane component	Spore maturation protein B
Phosphoglucosamine mutase (EC 5.4.2.10)	Hydroxymethylpyrimidine phosphate synthase ThiC (EC 4.1.99.17)	Hydroxymethylpyrimidine phosphate synthase ThiC (EC 4.1.99.17)	SSU ribosomal protein S10p (S20e)
Phosphoglycerate kinase (EC 2.7.2.3)	Hypothetical nudix hydrolase YeaB	Hypothetical nudix hydrolase YeaB	SSU ribosomal protein S11e (S17p)
Phosphonate ABC transporter permease protein phnE (TC 3.A.1.9.1)	Hypothetical protein (associated with DNA helicase - Rad25 homolog)	Hypothetical protein (associated with DNA helicase - Rad25 homolog)	SSU ribosomal protein S11p (S14e)
Phosphonate ABC transporter permease protein phnE1 (TC 3.A.1.9.1)	Hypothetical radical SAM family enzyme in heat shock gene cluster, similarity with CPO of BS HemN-type	Hypothetical radical SAM family enzyme in heat shock gene cluster, similarity with CPO of BS HemN-type	SSU ribosomal protein S13e (S15p)
Phosphopantetheine adenylyltransferase (EC 2.7.7.3)	Hypothetical transmembrane protein coupled to NADH-ubiquinone oxidoreductase chain 5 homolog	Hypothetical transmembrane protein coupled to NADH-ubiquinone oxidoreductase chain 5 homolog	SSU ribosomal protein S14e (S11p)
Phosphopantothenoylcysteine decarboxylase (EC 4.1.1.36)	Hypoxanthine/guanine permease PbuG	Hypoxanthine/guanine permease PbuG	SSU ribosomal protein S15Ae (S8p)
Phosphopantothenoylcysteine synthetase (EC 6.3.2.5)	Imidazole glycerol phosphate synthase amidotransferase	Imidazole glycerol phosphate synthase amidotransferase subunit (EC	SSU ribosomal protein S15e (S19p)

	subunit (EC 2.4.2.-)	2.4.2.-)	
Phosphoribosyl-AMP cyclohydrolase (EC 3.5.4.19)	**Imidazole glycerol phosphate synthase cyclase subunit (EC 4.1.3.-)**	**Imidazole glycerol phosphate synthase cyclase subunit (EC 4.1.3.-)**	SSU ribosomal protein S16e (S9p)
Phosphoribosylamine--glycine ligase (EC 6.3.4.13)	Imidazoleglycerol-phosphate dehydratase (EC 4.2.1.19)	Imidazoleglycerol-phosphate dehydratase (EC 4.2.1.19)	SSU ribosomal protein S17p (S11e)
Phosphoribosylaminoimidazole carboxylase catalytic subunit (EC 4.1.1.21)	Imidazolonepropionase (EC 3.5.2.7)	Imidazolonepropionase (EC 3.5.2.7)	SSU ribosomal protein S18e (S13p)
Phosphoribosylaminoimidazole-succinocarboxamide synthase (EC 6.3.2.6)	IMP cyclohydrolase (EC 3.5.4.10)	IMP cyclohydrolase (EC 3.5.4.10)	SSU ribosomal protein S19e
Phosphoribosylformimino-5-aminoimidazole carboxamide ribotide isomerase (EC 5.3.1.16)	IMP cyclohydrolase (EC 3.5.4.10) [alternate form]	IMP cyclohydrolase (EC 3.5.4.10) [alternate form]	SSU ribosomal protein S19p (S15e)
Phosphoribosylformylglycinamidine cyclo-ligase (EC 6.3.3.1)	Indole-3-glycerol phosphate synthase (EC 4.1.1.48)	Indole-3-glycerol phosphate synthase (EC 4.1.1.48)	SSU ribosomal protein S1p
Phosphoribosylformylglycinamidine synthase, glutamine amidotransferase subunit (EC 6.3.5.3)	Inorganic pyrophospatase PpaX (EC 3.1.3.18)	Inorganic pyrophospatase PpaX (EC 3.1.3.18)	SSU ribosomal protein S20e (S10p)
Phosphoribosylformylglycinamidine synthase, synthetase subunit (EC 6.3.5.3)	Inorganic pyrophosphatase (EC 3.6.1.1)	Inorganic pyrophosphatase (EC 3.6.1.1)	SSU ribosomal protein S23e (S12p)
Phosphosulfolactate synthase (EC 4.4.1.19)	Inosine-5'-monophosphate dehydrogenase (EC 1.1.1.205)	Inosine-5'-monophosphate dehydrogenase (EC 1.1.1.205)	SSU ribosomal protein S29e (S14p)
photosystem I P700 chlorophyll a apoprotein subunit Ia (PsaA)	Inosine-uridine preferring nucleoside hydrolase (EC 3.2.2.1)	Inosine-uridine preferring nucleoside hydrolase (EC 3.2.2.1)	SSU ribosomal protein S2e (S5p)
Photosystem II CP47 protein (PsbB)	Integration host factor beta subunit	Integration host factor beta subunit	SSU ribosomal protein S2p (SAe)
photosystem II protein D1 (PsbA)	intein-containing	intein-containing	SSU ribosomal protein S3Ae
photosystem II protein D2 (PsbD)	Iron-sulfur cluster assembly protein SufB	Iron-sulfur cluster assembly protein SufB	SSU ribosomal protein S3e (S3p)
Phycobilisome core-membrane linker polypeptide	Isocitrate dehydrogenase [NAD] (EC 1.1.1.41)	Isocitrate dehydrogenase [NAD] (EC 1.1.1.41)	SSU ribosomal protein S4e
Phycobilisome small core linker polypeptide	Isocitrate dehydrogenase [NADP] (EC 1.1.1.42)	Isocitrate dehydrogenase [NADP] (EC 1.1.1.42)	SSU ribosomal protein S5e (S7p)
Phytoene dehydrogenase (EC 1.14.99.-)	Isocitrate lyase (EC 4.1.3.1)	Isocitrate lyase (EC 4.1.3.1)	SSU ribosomal protein S8e
Phytoene desaturase (EC 1.14.99.-)	Isopentenyl-diphosphate delta-isomerase (EC 5.3.3.2)	Isopentenyl-diphosphate delta-isomerase (EC 5.3.3.2)	SSU ribosomal protein S9e (S4p)
Polyhydroxyalkanoic acid synthase	Kef-type K+ transport systems (NAD-binding component fused to domain related to exopolyphosphatase)	Kef-type K+ transport systems (NAD-binding component fused to domain related to exopolyphosphatase)	SSU ribosomal protein S9p (S16e)
Polyphosphate kinase (EC 2.7.4.1)	Ketol-acid reductoisomerase (EC 1.1.1.86)	Ketol-acid reductoisomerase (EC 1.1.1.86)	SSU ribosomal protein SAe (S2p)
Potassium channel protein	Ku domain protein	Ku domain protein	Succinate dehydrogenase cytochrome b subunit
Potassium efflux system KefA protein	Kynureninase (EC 3.7.1.3)	Kynureninase (EC 3.7.1.3)	Succinate dehydrogenase flavoprotein subunit (EC 1.3.99.1)
Potassium uptake protein TrkH	Kynurenine formamidase, bacterial (EC 3.5.1.9)	Kynurenine formamidase, bacterial (EC 3.5.1.9)	Succinate dehydrogenase hydrophobic membrane anchor protein

Potassium uptake protein, integral membrane component, KtrB	**L-arabinose transport system permease protein (TC 3.A.1.2.2)**	L-arabinose transport system permease protein (TC 3.A.1.2.2)	Succinate dehydrogenase iron-sulfur protein (EC 1.3.99.1)
Predicted N6-adenine-specific RNA methylase containing THUMP domain	L-aspartate oxidase (EC 1.4.3.16)	L-aspartate oxidase (EC 1.4.3.16)	Succinate-semialdehyde dehydrogenase [NADP+] (EC 1.2.1.79)
Predicted RNA-binding protein COG1491	L-lactate dehydrogenase (EC 1.1.1.27)	L-lactate dehydrogenase (EC 1.1.1.27)	Succinyl-CoA ligase [ADP-forming] alpha chain (EC 6.2.1.5)
Predicted transcriptional regulator of N-Acetylglucosamine utilization, GntR family	L-proline glycine betaine ABC transport system permease protein ProV (TC 3.A.1.12.1)	L-proline glycine betaine ABC transport system permease protein ProV (TC 3.A.1.12.1)	Succinyl-CoA ligase [ADP-forming] beta chain (EC 6.2.1.5)
Prephenate and/or arogenate dehydrogenase (unknown specificity) (EC 1.3.1.12)(EC 1.3.1.43)	L-proline glycine betaine ABC transport system permease protein ProW (TC 3.A.1.12.1)	L-proline glycine betaine ABC transport system permease protein ProW (TC 3.A.1.12.1)	Succinyl-CoA:3-ketoacid-coenzyme A transferase subunit B (EC 2.8.3.5)
Prephenate dehydratase (EC 4.2.1.51)	L-Proline/Glycine betaine transporter ProP	L-Proline/Glycine betaine transporter ProP	Sucrose-6-phosphate hydrolase (EC 3.2.1.B3)
Probable low-affinity inorganic phosphate transporter	L-serine dehydratase, alpha subunit (EC 4.3.1.17)	L-serine dehydratase, alpha subunit (EC 4.3.1.17)	Sulfate adenylyltransferase subunit 2 (EC 2.7.7.4)
Programmed cell death toxin YdcE	L-serine dehydratase, beta subunit (EC 4.3.1.17)	L-serine dehydratase, beta subunit (EC 4.3.1.17)	Sulfate permease
Propionyl-CoA carboxylase beta chain (EC 6.4.1.3)	Large-conductance mechanosensitive channel	Large-conductance mechanosensitive channel	Sulfur carrier protein adenylyltransferase ThiF
Proteasome subunit alpha (EC 3.4.25.1), archaeal	Lipid A export ATP-binding/permease protein MsbA	Lipid A export ATP-binding/permease protein MsbA	Superoxide dismutase [Mn] (EC 1.15.1.1)
Proteasome-activating AAA-ATPase (PAN), archaeal	Lipoate synthase	Lipoate synthase	Tagatose 1,6-bisphosphate aldolase (EC 4.1.2.40)
Protein YidD	Lipoate-protein ligase A	Lipoate-protein ligase A	TATA-box binding protein
Pterin-4-alpha-carbinolamine dehydratase (EC 4.2.1.96)	Lipoate-protein ligase A, C-terminal 70 percent	Lipoate-protein ligase A, C-terminal 70 percent	Thermosome subunit
Putative deoxyribonuclease YcfH	Lipoteichoic acid synthase LtaS Type IIIa	Lipoteichoic acid synthase LtaS Type IIIa	Thiaminase II (EC 3.5.99.2) involved in salvage of thiamin pyrimidine moiety
Putative formate dehydrogenase oxidoreductase protein	Long-chain-fatty-acid--CoA ligase (EC 6.2.1.3)	Long-chain-fatty-acid--CoA ligase (EC 6.2.1.3)	ThiJ/PfpI family protein
Putative periplasmic phosphate-binding protein PstS (Halobacteriales type)	Low G+C gram positive nudix hydrolase YtkD (EC 3.6.-.-)	Low G+C gram positive nudix hydrolase YtkD (EC 3.6.-.-)	Thioredoxin reductase (EC 1.8.1.9)
putative sodium-dependent bicarbonate transporter	low-specificity D-threonine aldolase	low-specificity D-threonine aldolase	Thiosulfate sulfurtransferase, rhodanese (EC 2.8.1.1)
Putative stomatin/prohibitin-family membrane protease subunit aq_911	Low-specificity L-threonine aldolase (EC 4.1.2.48)	Low-specificity L-threonine aldolase (EC 4.1.2.48)	Threonine dehydratase (EC 4.3.1.19)
Putative Zn-dependent hydrolase in polyisoprenoid biosynthetic cluster	LSU m3Psi1915 methyltransferase RlmH	LSU m3Psi1915 methyltransferase RlmH	Threonine dehydrogenase and related Zn-dependent dehydrogenases
Pyridoxamine 5'-phosphate oxidase (EC 1.4.3.5)	LSU ribosomal protein L10Ae (L1p)	LSU ribosomal protein L10Ae (L1p)	Threonine synthase (EC 4.2.3.1)
Pyridoxine 5'-phosphate synthase (EC 2.6.99.2)	LSU ribosomal protein L10p (P0)	LSU ribosomal protein L10p (P0)	Threonyl-tRNA synthetase (EC 6.1.1.3)
Pyridoxine biosynthesis glutamine amidotransferase, glutaminase subunit (EC 2.4.2.-)	LSU ribosomal protein L11p (L12e)	LSU ribosomal protein L11p (L12e)	Thymidine kinase (EC 2.7.1.21)
Pyridoxine biosynthesis glutamine amidotransferase,	LSU ribosomal protein L12e (L11p)	LSU ribosomal protein L12e (L11p)	Thymidylate kinase (EC 2.7.4.9)

synthase subunit (EC 2.4.2.-)			
Pyrimidine-nucleoside phosphorylase (EC 2.4.2.2)	LSU ribosomal protein L15e	LSU ribosomal protein L15e	Thymidylate synthase (EC 2.1.1.45)
Pyrophosphate-energized proton pump (EC 3.6.1.1)	LSU ribosomal protein L16p (L10e)	LSU ribosomal protein L16p (L10e)	Thymidylate synthase thyX (EC 2.1.1.-)
Pyrroline-5-carboxylate reductase (EC 1.5.1.2)	LSU ribosomal protein L17e (L22p)	LSU ribosomal protein L17e (L22p)	tolB protein precursor, periplasmic protein involved in the tonb-independent uptake of group A colicins
Pyruvate dehydrogenase E1 component alpha subunit (EC 1.2.4.1)	LSU ribosomal protein L18e	LSU ribosomal protein L18e	TonB-dependent receptor
Pyruvate dehydrogenase E1 component beta subunit (EC 1.2.4.1)	LSU ribosomal protein L18p (L5e)	LSU ribosomal protein L18p (L5e)	Topoisomerase IV subunit A (EC 5.99.1.-)
Pyruvate kinase (EC 2.7.1.40)	LSU ribosomal protein L1e (L4p)	LSU ribosomal protein L1e (L4p)	Topoisomerase IV subunit B (EC 5.99.1.-)
Pyruvate-flavodoxin oxidoreductase (EC 1.2.7.-)	LSU ribosomal protein L1p (L10Ae)	LSU ribosomal protein L1p (L10Ae)	Transaldolase (EC 2.2.1.2)
Pyruvate,phosphate dikinase (EC 2.7.9.1)	LSU ribosomal protein L20p	LSU ribosomal protein L20p	Transcription antitermination protein NusG
Pyruvate:ferredoxin oxidoreductase, beta subunit (EC 1.2.7.1)	LSU ribosomal protein L23Ae (L23p)	LSU ribosomal protein L23Ae (L23p)	Transcription initiation factor B
Pyruvate:ferredoxin oxidoreductase, delta subunit (EC 1.2.7.1)	LSU ribosomal protein L23e (L14p)	LSU ribosomal protein L23e (L14p)	Transcription termination factor Rho
Queuosine Biosynthesis QueC ATPase	LSU ribosomal protein L24p (L26e)	LSU ribosomal protein L24p (L26e)	Transcription termination protein NusA
Queuosine biosynthesis QueD, PTPS-I	LSU ribosomal protein L2p (L8e)	LSU ribosomal protein L2p (L8e)	Transcription-repair coupling factor
Queuosine Biosynthesis QueE Radical SAM	LSU ribosomal protein L37e	LSU ribosomal protein L37e	Transketolase, N-terminal section (EC 2.2.1.1)
Radical SAM domain heme biosynthesis protein	LSU ribosomal protein L3e (L3p)	LSU ribosomal protein L3e (L3p)	Translation elongation factor 1 alpha subunit
RecA protein	LSU ribosomal protein L4p (L1e)	LSU ribosomal protein L4p (L1e)	Translation elongation factor 1 beta subunit
RecD-like DNA helicase YrrC	LSU ribosomal protein L5p (L11e)	LSU ribosomal protein L5p (L11e)	Translation elongation factor 2
Recombination inhibitory protein MutS2	LSU ribosomal protein L6p (L9e)	LSU ribosomal protein L6p (L9e)	Translation elongation factor G
Recombination protein RecR	LSU ribosomal protein L7/L12 (P1/P2)	LSU ribosomal protein L7/L12 (P1/P2)	Translation elongation factor G-related protein
Replication factor A (ssDNA-binding protein)	LSU ribosomal protein L7Ae	LSU ribosomal protein L7Ae	Translation elongation factor Tu
Replication factor C large subunit	LSU ribosomal protein L7e (L30p)	LSU ribosomal protein L7e (L30p)	Translation initiation factor 1
Replication factor C small subunit	LSU ribosomal protein L8e (L2p)	LSU ribosomal protein L8e (L2p)	Translation initiation factor 2
Replicative DNA helicase (DnaB) (EC 3.6.4.12)	LSU ribosomal protein L9e (L6p)	LSU ribosomal protein L9e (L6p)	Translation initiation factor 2B alpha/beta/delta-type subunit
Respiratory nitrate reductase alpha chain (EC 1.7.99.4)	LSU ribosomal protein P0 (L10p)	LSU ribosomal protein P0 (L10p)	Translation initiation factor SUI1-related protein
Riboflavin synthase eubacterial/eukaryotic (EC 2.5.1.9)	Lycopene elongase (EC 2.5.1.-)	Lycopene elongase (EC 2.5.1.-)	Trehalose synthase (EC 5.4.99.16)
Ribokinase (EC 2.7.1.15)	Lysine 2,3-aminomutase (EC 5.4.3.2)	Lysine 2,3-aminomutase (EC 5.4.3.2)	Triosephosphate isomerase (EC 5.3.1.1)

Ribonuclease BN (EC 3.1.-.-)	Lysine decarboxylase (EC 4.1.1.18)	Lysine decarboxylase (EC 4.1.1.18)	Trk system potassium uptake protein TrkA
Ribonuclease HII (EC 3.1.26.4)	Lysyl-tRNA synthetase (class I) (EC 6.1.1.6)	Lysyl-tRNA synthetase (class I) (EC 6.1.1.6)	tRNA (5-methylaminomethyl-2-thiouridylate)-methyltransferase (EC 2.1.1.61)
Ribonucleotide reductase of class Ia (aerobic), alpha subunit (EC 1.17.4.1)	Lysyl-tRNA synthetase (class II) (EC 6.1.1.6)	Lysyl-tRNA synthetase (class II) (EC 6.1.1.6)	tRNA (Guanine37-N1) -methyltransferase (EC 2.1.1.31)
Ribonucleotide reductase of class II (coenzyme B12-dependent) (EC 1.17.4.1)	Macrolide export ATP-binding/permease protein MacB (EC 3.6.3.-)	Macrolide export ATP-binding/permease protein MacB (EC 3.6.3.-)	tRNA methylase Trm12p Wyeosine biosynthesis
Ribose ABC transport system, permease protein RbsC (TC 3.A.1.2.1)	Magnesium and cobalt efflux protein CorC	Magnesium and cobalt efflux protein CorC	tRNA nucleotidyltransferase (EC 2.7.7.21) (EC 2.7.7.25)
Ribose-phosphate pyrophosphokinase (EC 2.7.6.1)	Malate dehydrogenase (EC 1.1.1.37)	Malate dehydrogenase (EC 1.1.1.37)	tRNA nucleotidyltransferase, archaeal type (EC 2.7.7.21) (EC 2.7.7.25)
Ribosome small subunit-stimulated GTPase EngC	Malate Na(+) symporter	Malate Na(+) symporter	tRNA pseudouridine 55 synthase (EC 4.2.1.70)
Ribulose bisphosphate carboxylase large chain (EC 4.1.1.39)	Malate:quinone oxidoreductase (EC 1.1.5.4)	Malate:quinone oxidoreductase (EC 1.1.5.4)	tRNA pseudouridine synthase A (EC 4.2.1.70)
Ribulose-phosphate 3-epimerase (EC 5.1.3.1)	Malonyl CoA-acyl carrier protein transacylase (EC 2.3.1.39)	Malonyl CoA-acyl carrier protein transacylase (EC 2.3.1.39)	tRNA-i(6)A37 methylthiotransferase
RNA polymerase sigma factor for flagellar operon	Malto-oligosyltrehalose trehalohydrolase (EC 3.2.1.141)	Malto-oligosyltrehalose trehalohydrolase (EC 3.2.1.141)	tRNA-intron endonuclease (EC 3.1.27.9)
RNA polymerase sigma-54 factor RpoN	Maltodextrin glucosidase (EC 3.2.1.20)	Maltodextrin glucosidase (EC 3.2.1.20)	Tryptophan synthase alpha chain (EC 4.2.1.20)
RND efflux system, inner membrane transporter CmeB	Maltose/maltodextrin ABC transporter, permease protein MalF	Maltose/maltodextrin ABC transporter, permease protein MalF	Tryptophan synthase beta chain (EC 4.2.1.20)
Rod shape-determining protein MreB	Maltose/maltodextrin ABC transporter, permease protein MalG	Maltose/maltodextrin ABC transporter, permease protein MalG	Tryptophanase (EC 4.1.99.1)
Rod shape-determining protein MreC	Maltose/maltodextrin ABC transporter, substrate binding periplasmic protein MalE	Maltose/maltodextrin ABC transporter, substrate binding periplasmic protein MalE	Tryptophanyl-tRNA synthetase (EC 6.1.1.2)
Rod shape-determining protein RodA	Maltose/maltodextrin transport ATP-binding protein MalK (EC 3.6.3.19)	Maltose/maltodextrin transport ATP-binding protein MalK (EC 3.6.3.19)	TsaD/Kae1/Qri7 protein, required for threonylcarbamoyladenosine t(6)A37 formation in tRNA
Rossmann fold nucleotide-binding protein Smf possibly involved in DNA uptake	Manganese superoxide dismutase (EC 1.15.1.1)	Manganese superoxide dismutase (EC 1.15.1.1)	Type I restriction-modification system, restriction subunit R (EC 3.1.21.3)
rRNA biogenesis protein Nop5/Nop56	Manganese-dependent inorganic pyrophosphatase (EC 3.6.1.1)	Manganese-dependent inorganic pyrophosphatase (EC 3.6.1.1)	Type II secretory pathway, ATPase PulE/Tfp pilus assembly pathway, ATPase PilB
rRNA small subunit methyltransferase H	Mannitol operon activator, BglG family	Mannitol operon activator, BglG family	Tyrosyl-tRNA synthetase (EC 6.1.1.1)
rRNA small subunit methyltransferase I	Mannose-1-phosphate guanylyltransferase (EC 2.7.7.13)	Mannose-1-phosphate guanylyltransferase (EC 2.7.7.13)	Ubiquinol--cytochrome c reductase, cytochrome B subunit (EC 1.10.2.2)
S-adenosylmethionine synthetase (EC 2.5.1.6)	Membrane-associated zinc metalloprotease	Membrane-associated zinc metalloprotease	**Ubiquinone/menaquinone biosynthesis methyltransferase UbiE (EC 2.1.1.-)**
S-adenosylmethionine:tRNA ribosyltransferase-isomerase (EC 5.-.-.-)	Membrane-attached cytochrome c550	Membrane-attached cytochrome c550	UDP-glucose 4-epimerase (EC 5.1.3.2)
Segregation and condensation protein B	Menaquinone-specific isochorismate synthase (EC 5.4.4.2)	Menaquinone-specific isochorismate synthase (EC 5.4.4.2)	UDP-N-acetylmuramoylalanyl-D-glutamyl-2,6-diaminopimelate--D-

			alanyl-D-alanine ligase (EC 6.3.2.10)
Serine hydroxymethyltransferase (EC 2.1.2.1)	Mercuric ion reductase (EC 1.16.1.1)	Mercuric ion reductase (EC 1.16.1.1)	Universal stress protein family
Serine phosphatase RsbU, regulator of sigma subunit	Methenyltetrahydrofolate cyclohydrolase (EC 3.5.4.9)	Methenyltetrahydrofolate cyclohydrolase (EC 3.5.4.9)	UPF0129 protein MJ1474
Seryl-tRNA synthetase (EC 6.1.1.11)	Methionine ABC transporter ATP-binding protein	Methionine ABC transporter ATP-binding protein	Uracil-DNA glycosylase, family 4
Signal peptidase I (EC 3.4.21.89)	Methionine ABC transporter permease protein	Methionine ABC transporter permease protein	Urease alpha subunit (EC 3.5.1.5)
Signal peptidase, type IV - prepilin/preflagellin	Methionine ABC transporter substrate-binding protein	Methionine ABC transporter substrate-binding protein	Urocanate hydratase (EC 4.2.1.49)
Signal recognition particle receptor protein FtsY (=alpha subunit) (TC 3.A.5.1.1)	Methionine gamma-lyase (EC 4.4.1.11)	Methionine gamma-lyase (EC 4.4.1.11)	Uroporphyrinogen-III methyltransferase (EC 2.1.1.107)
Signal recognition particle, subunit Ffh SRP54 (TC 3.A.5.1.1)	Methionyl-tRNA formyltransferase (EC 2.1.2.9)	Methionyl-tRNA formyltransferase (EC 2.1.2.9)	V-type ATP synthase subunit A (EC 3.6.3.14)
Similar to Hydroxyacylglutathione hydrolase, but in an organism lacking glutathione biosynthesis	Methylcrotonyl-CoA carboxylase carboxyl transferase subunit (EC 6.4.1.4)	Methylcrotonyl-CoA carboxylase carboxyl transferase subunit (EC 6.4.1.4)	V-type ATP synthase subunit B (EC 3.6.3.14)
Single-stranded DNA-binding protein	Methylglyoxal synthase (EC 4.2.3.3)	Methylglyoxal synthase (EC 4.2.3.3)	V-type ATP synthase subunit D (EC 3.6.3.14)
Solanesyl diphosphate synthase (EC 2.5.1.11)	Methylmalonate-semialdehyde dehydrogenase (EC 1.2.1.27)	Methylmalonate-semialdehyde dehydrogenase (EC 1.2.1.27)	V-type ATP synthase subunit I (EC 3.6.3.14)
SOS-response repressor and protease LexA (EC 3.4.21.88)	Methylmalonyl-CoA mutase (EC 5.4.99.2)	Methylmalonyl-CoA mutase (EC 5.4.99.2)	Various polyols ABC transporter, periplasmic substrate-binding protein
Spermidine Putrescine ABC transporter permease component PotB (TC 3.A.1.11.1)	Methylthioribose-1-phosphate isomerase (EC 5.3.1.23)	Methylthioribose-1-phosphate isomerase (EC 5.3.1.23)	Vitamin B12 ABC transporter, permease component BtuC
Spore maturation protein B	Molybdenum cofactor biosynthesis protein MoaB	Molybdenum cofactor biosynthesis protein MoaB	Xanthine-guanine phosphoribosyltransferase (EC 2.4.2.22)
SSU ribosomal protein S10p (S20e)	Molybdopterin biosynthesis protein MoeA	Molybdopterin biosynthesis protein MoeA	Xanthine/uracil/thiamine/ascorbate permease family protein
SSU ribosomal protein S14e (S11p)	Molybdopterin-guanine dinucleotide biosynthesis protein MobB	Molybdopterin-guanine dinucleotide biosynthesis protein MobB	Xylose isomerase (EC 5.3.1.5)
SSU ribosomal protein S15Ae (S8p)	Monoamine oxidase (1.4.3.4)	Monoamine oxidase (1.4.3.4)	YbbM seven transmembrane helix protein
SSU ribosomal protein S16e (S9p)	Monoglyceride lipase (EC 3.1.1.23)	Monoglyceride lipase (EC 3.1.1.23)	Zinc ABC transporter, inner membrane permease protein ZnuB
SSU ribosomal protein S17p (S11e)	Multi antimicrobial extrusion protein (Na(+)/drug antiporter), MATE family of MDR efflux pumps	Multi antimicrobial extrusion protein (Na(+)/drug antiporter), MATE family of MDR efflux pumps	
SSU ribosomal protein S1p	Multicopper oxidase	Multicopper oxidase	
SSU ribosomal protein S23e (S12p)	Multimodular transpeptidase-transglycosylase (EC 2.4.1.129) (EC 3.4.-.-)	Multimodular transpeptidase-transglycosylase (EC 2.4.1.129) (EC 3.4.-.-)	
SSU ribosomal protein S24e	Multiple sugar ABC transporter, ATP-binding protein	Multiple sugar ABC transporter, ATP-binding protein	
SSU ribosomal protein S2e (S5p)	N-acetyl-gamma-glutamyl-phosphate reductase (EC	N-acetyl-gamma-glutamyl-phosphate reductase (EC 1.2.1.38)	

	1.2.1.38)	
SSU ribosomal protein S2p (SAe)	N-acetyl-L,L-diaminopimelate aminotransferase (EC 2.6.1.-)	N-acetyl-L,L-diaminopimelate aminotransferase (EC 2.6.1.-)
SSU ribosomal protein S3e (S3p)	N-acetyl-L,L-diaminopimelate deacetylase (EC 3.5.1.47)	N-acetyl-L,L-diaminopimelate deacetylase (EC 3.5.1.47)
SSU ribosomal protein S3p (S3e)	N-acetylglucosamine-1-phosphate uridyltransferase (EC 2.7.7.23)	N-acetylglucosamine-1-phosphate uridyltransferase (EC 2.7.7.23)
SSU ribosomal protein S4e	N-acetylglucosamine-6-phosphate deacetylase (EC 3.5.1.25)	N-acetylglucosamine-6-phosphate deacetylase (EC 3.5.1.25)
SSU ribosomal protein S4p (S9e)	N-acetylglutamate synthase (EC 2.3.1.1)	N-acetylglutamate synthase (EC 2.3.1.1)
SSU ribosomal protein S5e (S7p)	N-acetylmuramic acid 6-phosphate etherase	N-acetylmuramic acid 6-phosphate etherase
SSU ribosomal protein S5p (S2e)	N-acetylneuraminate synthase (EC 2.5.1.56)	N-acetylneuraminate synthase (EC 2.5.1.56)
SSU ribosomal protein S8e	N-acyl homoserine lactone hydrolase	N-acyl homoserine lactone hydrolase
SSU ribosomal protein S8p (S15Ae)	N-methylhydantoinase (ATP-hydrolyzing) (EC 3.5.2.14)	N-methylhydantoinase (ATP-hydrolyzing) (EC 3.5.2.14)
SSU ribosomal protein S9e (S4p)	N-methylhydantoinase B (EC 3.5.2.14)	N-methylhydantoinase B (EC 3.5.2.14)
Succinate dehydrogenase cytochrome b-556 subunit	N(5),N(10)-methenyltetrahydromethanopterin cyclohydrolase (EC 3.5.4.27)	N(5),N(10)-methenyltetrahydromethanopterin cyclohydrolase (EC 3.5.4.27)
Succinate dehydrogenase flavoprotein subunit (EC 1.3.99.1)	Na(+) H(+) antiporter subunit D	Na(+) H(+) antiporter subunit D
Succinate dehydrogenase iron-sulfur protein (EC 1.3.99.1)	NAD synthetase (EC 6.3.1.5)	NAD synthetase (EC 6.3.1.5)
Succinate-semialdehyde dehydrogenase [NADP+] (EC 1.2.1.79)	NAD-dependent formate dehydrogenase alpha subunit	NAD-dependent formate dehydrogenase alpha subunit
Succinyl-CoA ligase [ADP-forming] beta chain (EC 6.2.1.5)	NAD-dependent formate dehydrogenase beta subunit	NAD-dependent formate dehydrogenase beta subunit
Sulfate adenylyltransferase subunit 2 (EC 2.7.7.4)	NAD-dependent glyceraldehyde-3-phosphate dehydrogenase (EC 1.2.1.12)	NAD-dependent glyceraldehyde-3-phosphate dehydrogenase (EC 1.2.1.12)
Sulfate and thiosulfate import ATP-binding protein CysA (EC 3.6.3.25)	NAD-dependent malic enzyme (EC 1.1.1.38)	NAD-dependent malic enzyme (EC 1.1.1.38)
Sulfate permease	NAD-dependent protein deacetylase of SIR2 family	NAD-dependent protein deacetylase of SIR2 family
Sulfite reductase [NADPH] hemoprotein beta-component (EC 1.8.1.2)	NAD-independent protein deacetylase AcuC	NAD-independent protein deacetylase AcuC
Sulfur carrier protein adenylyltransferase ThiF	NAD-reducing hydrogenase subunit HoxF (EC 1.12.1.2)	NAD-reducing hydrogenase subunit HoxF (EC 1.12.1.2)
Superoxide dismutase [Mn] (EC 1.15.1.1)	NAD-specific glutamate dehydrogenase (EC 1.4.1.2)	NAD-specific glutamate dehydrogenase (EC 1.4.1.2)
TATA-box binding protein	NAD(P) transhydrogenase alpha subunit (EC 1.6.1.2)	NAD(P) transhydrogenase alpha subunit (EC 1.6.1.2)
Thermosome subunit	NAD(P) transhydrogenase subunit beta (EC 1.6.1.2)	NAD(P) transhydrogenase subunit beta (EC 1.6.1.2)

Thiaminase II (EC 3.5.99.2) involved in salvage of thiamin pyrimidine moiety	NAD(P)-dependent glyceraldehyde 3-phosphate dehydrogenase archaeal (EC 1.2.1.59)	NAD(P)-dependent glyceraldehyde 3-phosphate dehydrogenase archaeal (EC 1.2.1.59)
Thiamine-monophosphate kinase (EC 2.7.4.16)	NAD(P)HX epimerase	NAD(P)HX epimerase
Thiazole biosynthesis protein ThiG	NADH dehydrogenase (EC 1.6.99.3)	NADH dehydrogenase (EC 1.6.99.3)
Thioredoxin	NADH dehydrogenase, subunit 5	NADH dehydrogenase, subunit 5
Thioredoxin reductase (EC 1.8.1.9)	NADH-dependent butanol dehydrogenase A (EC 1.1.1.-)	NADH-dependent butanol dehydrogenase A (EC 1.1.1.-)
Threonine dehydratase (EC 4.3.1.19)	NADH-ubiquinone oxidoreductase chain B (EC 1.6.5.3)	NADH-ubiquinone oxidoreductase chain B (EC 1.6.5.3)
Threonine dehydratase, catabolic (EC 4.3.1.19)	NADH-ubiquinone oxidoreductase chain C (EC 1.6.5.3)	NADH-ubiquinone oxidoreductase chain C (EC 1.6.5.3)
Threonine synthase (EC 4.2.3.1)	NADH-ubiquinone oxidoreductase chain D (EC 1.6.5.3)	NADH-ubiquinone oxidoreductase chain D (EC 1.6.5.3)
Threonyl-tRNA synthetase (EC 6.1.1.3)	NADH-ubiquinone oxidoreductase chain E (EC 1.6.5.3)	NADH-ubiquinone oxidoreductase chain E (EC 1.6.5.3)
Thymidylate kinase (EC 2.7.4.9)	NADH-ubiquinone oxidoreductase chain F (EC 1.6.5.3)	NADH-ubiquinone oxidoreductase chain F (EC 1.6.5.3)
tmRNA-binding protein SmpB	NADH-ubiquinone oxidoreductase chain G (EC 1.6.5.3)	NADH-ubiquinone oxidoreductase chain G (EC 1.6.5.3)
tolB protein precursor, periplasmic protein involved in the tonb-independent uptake of group A colicins	NADH-ubiquinone oxidoreductase chain H (EC 1.6.5.3)	NADH-ubiquinone oxidoreductase chain H (EC 1.6.5.3)
TonB-dependent receptor	NADH-ubiquinone oxidoreductase chain I (EC 1.6.5.3)	NADH-ubiquinone oxidoreductase chain I (EC 1.6.5.3)
Topoisomerase IV subunit A (EC 5.99.1.-)	NADH-ubiquinone oxidoreductase chain K (EC 1.6.5.3)	NADH-ubiquinone oxidoreductase chain K (EC 1.6.5.3)
Topoisomerase IV subunit B (EC 5.99.1.-)	NADH-ubiquinone oxidoreductase chain L (EC 1.6.5.3)	NADH-ubiquinone oxidoreductase chain L (EC 1.6.5.3)
Transaldolase (EC 2.2.1.2)	NADH-ubiquinone oxidoreductase chain M (EC 1.6.5.3)	NADH-ubiquinone oxidoreductase chain M (EC 1.6.5.3)
Transcription antitermination protein NusG	NADH-ubiquinone oxidoreductase chain N (EC 1.6.5.3)	NADH-ubiquinone oxidoreductase chain N (EC 1.6.5.3)
Transcription initiation factor B	NADP-dependent malic enzyme (EC 1.1.1.40)	NADP-dependent malic enzyme (EC 1.1.1.40)
Transcription termination factor Rho	NADP-specific glutamate dehydrogenase (EC 1.4.1.4)	NADP-specific glutamate dehydrogenase (EC 1.4.1.4)
Transcription termination protein NusA	NADPH-dependent glyceraldehyde-3-phosphate dehydrogenase (EC 1.2.1.13)	NADPH-dependent glyceraldehyde-3-phosphate dehydrogenase (EC 1.2.1.13)
Transcription-repair coupling factor	Naphthoate synthase (EC 4.1.3.36)	Naphthoate synthase (EC 4.1.3.36)

transcriptional regulator, Crp/Fnr family	Neopullulanase (EC 3.2.1.135)	Neopullulanase (EC 3.2.1.135)
Transketolase, N-terminal section (EC 2.2.1.1)	NG,NG-dimethylarginine dimethylaminohydrolase 1 (EC 3.5.3.18)	NG,NG-dimethylarginine dimethylaminohydrolase 1 (EC 3.5.3.18)
Translation elongation factor 1 alpha subunit	Nicotinamidase (EC 3.5.1.19)	Nicotinamidase (EC 3.5.1.19)
Translation elongation factor 2	Nicotinamidase family protein YcaC	Nicotinamidase family protein YcaC
Translation elongation factor G	Nicotinamide-nucleotide adenylyltransferase, NadM family (EC 2.7.7.1)	Nicotinamide-nucleotide adenylyltransferase, NadM family (EC 2.7.7.1)
Translation elongation factor G-related protein	Nicotinate phosphoribosyltransferase (EC 2.4.2.11)	Nicotinate phosphoribosyltransferase (EC 2.4.2.11)
Translation elongation factor LepA	Nicotinate-nucleotide adenylyltransferase (EC 2.7.7.18)	Nicotinate-nucleotide adenylyltransferase (EC 2.7.7.18)
Translation elongation factor P	NifS-like cysteine desulfurase DndA	NifS-like cysteine desulfurase DndA
Translation elongation factor Ts	Nitric oxide reductase activation protein NorQ	Nitric oxide reductase activation protein NorQ
Translation elongation factor Tu	Nitric oxide synthase oxygenase (EC 1.-.-.-)	Nitric oxide synthase oxygenase (EC 1.-.-.-)
Translation initiation factor 1A	Nitric-oxide reductase (EC 1.7.99.7), quinol-dependent	Nitric-oxide reductase (EC 1.7.99.7), quinol-dependent
Translation initiation factor 2	Nitrite reductase probable [NAD(P)H] subunit (EC 1.7.1.4)	Nitrite reductase probable [NAD(P)H] subunit (EC 1.7.1.4)
Translation initiation factor SUI1-related protein	Nitrogen regulatory protein P-II	Nitrogen regulatory protein P-II
Trehalose synthase (EC 5.4.99.16)	Nitrogenase FeMo-cofactor scaffold and assembly protein NifN	Nitrogenase FeMo-cofactor scaffold and assembly protein NifN
Trehalose-6-phosphate hydrolase (EC 3.2.1.93)	Nitrous oxide reductase maturation protein NosF (ATPase)	Nitrous oxide reductase maturation protein NosF (ATPase)
Trehalose-6-phosphate phosphatase (EC 3.1.3.12)	Nitrous-oxide reductase (EC 1.7.99.6)	Nitrous-oxide reductase (EC 1.7.99.6)
Triosephosphate isomerase (EC 5.3.1.1)	Non-specific DNA-binding protein Dps	Non-specific DNA-binding protein Dps
Trk system potassium uptake protein TrkA	Nucleoside diphosphate kinase (EC 2.7.4.6)	Nucleoside diphosphate kinase (EC 2.7.4.6)
tRNA (5-methylaminomethyl-2-thiouridylate)-methyltransferase (EC 2.1.1.61)	Nucleoside permease NupC	Nucleoside permease NupC
tRNA (Guanine37-N1) -methyltransferase (EC 2.1.1.31)	Nucleoside triphosphate pyrophosphohydrolase MazG (EC 3.6.1.8)	Nucleoside triphosphate pyrophosphohydrolase MazG (EC 3.6.1.8)
tRNA nucleotidyltransferase (EC 2.7.7.21) (EC 2.7.7.25)	Nucleotidase YfbR, HD superfamily	Nucleotidase YfbR, HD superfamily
tRNA pseudouridine synthase A (EC 4.2.1.70)	O-acetylhomoserine sulfhydrylase (EC 2.5.1.49)	O-acetylhomoserine sulfhydrylase (EC 2.5.1.49)
tRNA-i(6)A37 methylthiotransferase	O-succinylbenzoate synthase (EC 4.2.1.113)	O-succinylbenzoate synthase (EC 4.2.1.113)
Tryptophan synthase alpha chain (EC 4.2.1.20)	O-succinylhomoserine sulfhydrylase (EC 2.5.1.48)	O-succinylhomoserine sulfhydrylase (EC 2.5.1.48)
Tryptophan synthase beta chain (EC 4.2.1.20)	Oligo-1,6-glucosidase (EC 3.2.1.10)	Oligo-1,6-glucosidase (EC 3.2.1.10)

Tryptophanyl-tRNA synthetase (EC 6.1.1.2)	Oligopeptidase A (EC 3.4.24.70)	Oligopeptidase A (EC 3.4.24.70)
TsaC protein (YrdC domain) required for threonylcarbamoyladenosine t(6)A37 modification in tRNA	Oligopeptide ABC transporter, periplasmic oligopeptide-binding protein OppA (TC 3.A.1.5.1)	Oligopeptide ABC transporter, periplasmic oligopeptide-binding protein OppA (TC 3.A.1.5.1)
Type I restriction-modification system, DNA-methyltransferase subunit M (EC 2.1.1.72)	Oligopeptide transport ATP-binding protein OppF (TC 3.A.1.5.1)	Oligopeptide transport ATP-binding protein OppF (TC 3.A.1.5.1)
Type I restriction-modification system, specificity subunit S (EC 3.1.21.3)	Oligopeptide transport system permease protein OppC (TC 3.A.1.5.1)	Oligopeptide transport system permease protein OppC (TC 3.A.1.5.1)
Type II secretory pathway, ATPase PulE/Tfp pilus assembly pathway, ATPase PilB	Origin of replication recognition protein	Origin of replication recognition protein
Type III restriction-modification system methylation subunit (EC 2.1.1.72)	Ornithine aminotransferase (EC 2.6.1.13)	Ornithine aminotransferase (EC 2.6.1.13)
Tyrosyl-tRNA synthetase (EC 6.1.1.1)	Ornithine carbamoyltransferase (EC 2.1.3.3)	Ornithine carbamoyltransferase (EC 2.1.3.3)
Ubiquinone biosynthesis monooxygenase UbiB	Ornithine cyclodeaminase (EC 4.3.1.12)	Ornithine cyclodeaminase (EC 4.3.1.12)
Ubiquinone/menaquinone biosynthesis methyltransferase UbiE (EC 2.1.1.-)	Orotate phosphoribosyltransferase (EC 2.4.2.10)	Orotate phosphoribosyltransferase (EC 2.4.2.10)
UDP-glucose 4-epimerase (EC 5.1.3.2)	Orotidine 5'-phosphate decarboxylase (EC 4.1.1.23)	Orotidine 5'-phosphate decarboxylase (EC 4.1.1.23)
UDP-N-acetylenolpyruvoylglucosamine reductase (EC 1.1.1.158)	oxidoreductase of aldo/keto reductase family, subgroup 1	oxidoreductase of aldo/keto reductase family, subgroup 1
UDP-N-acetylglucosamine 2-epimerase (EC 5.1.3.14)	Pantoate kinase, archaeal (EC 2.7.1.-)	Pantoate kinase, archaeal (EC 2.7.1.-)
UDP-N-acetylmuramoylalanyl-D-glutamyl-2,6-diaminopimelate--D-alanyl-D-alanine ligase (EC 6.3.2.10)	Pantoate--beta-alanine ligase (EC 6.3.2.1)	Pantoate--beta-alanine ligase (EC 6.3.2.1)
Universal stress protein family	Pantothenate kinase type II, eukaryotic (EC 2.7.1.33)	Pantothenate kinase type II, eukaryotic (EC 2.7.1.33)
Uracil phosphoribosyltransferase (EC 2.4.2.9)	Pantothenate:Na+ symporter (TC 2.A.21.1.1)	Pantothenate:Na+ symporter (TC 2.A.21.1.1)
Uracil-DNA glycosylase, family 4	Para-aminobenzoate synthase, amidotransferase component (EC 2.6.1.85)	Para-aminobenzoate synthase, amidotransferase component (EC 2.6.1.85)
Urea ABC transporter, urea binding protein	**Para-aminobenzoate synthase, aminase component (EC 2.6.1.85)**	**Para-aminobenzoate synthase, aminase component (EC 2.6.1.85)**
Urease alpha subunit (EC 3.5.1.5)	Paraquat-inducible protein B	Paraquat-inducible protein B
V-type ATP synthase subunit A (EC 3.6.3.14)	Penicillin-binding protein 1A/1B (PBP1)	Penicillin-binding protein 1A/1B (PBP1)
V-type ATP synthase subunit B (EC 3.6.3.14)	Penicillin-binding protein 2 (PBP-2)	Penicillin-binding protein 2 (PBP-2)
V-type ATP synthase subunit D (EC 3.6.3.14)	Peptide chain release factor 1	Peptide chain release factor 1
V-type ATP synthase subunit F (EC 3.6.3.14)	Peptide chain release factor 2	Peptide chain release factor 2

V-type ATP synthase subunit I (EC 3.6.3.14)	Peptide deformylase (EC 3.5.1.88)	Peptide deformylase (EC 3.5.1.88)
Various polyols ABC transporter, periplasmic substrate-binding protein	Peptidyl-prolyl cis-trans isomerase (EC 5.2.1.8)	Peptidyl-prolyl cis-trans isomerase (EC 5.2.1.8)
Various polyols ABC transporter, permease component 1	Peptidyl-prolyl cis-trans isomerase ppiA precursor (EC 5.2.1.8)	Peptidyl-prolyl cis-trans isomerase ppiA precursor (EC 5.2.1.8)
Vitamin B12 ABC transporter, B12-binding component BtuF	Permease of the drug/metabolite transporter (DMT) superfamily	Permease of the drug/metabolite transporter (DMT) superfamily
Vitamin B12 ABC transporter, permease component BtuC	Peroxide stress regulator PerR, FUR family	Peroxide stress regulator PerR, FUR family
Xanthine-guanine phosphoribosyltransferase (EC 2.4.2.22)	Phage integrase	Phage integrase
YbbL ABC transporter ATP-binding protein	Phage portal protein	Phage portal protein
YbbM seven transmembrane helix protein	Phage tail assembly protein	Phage tail assembly protein
Zinc ABC transporter, inner membrane permease protein ZnuB	Phage tail fiber protein	Phage tail fiber protein
	Phenylalanine-4-hydroxylase (EC 1.14.16.1)	Phenylalanine-4-hydroxylase (EC 1.14.16.1)
	Phenylalanyl-tRNA synthetase alpha chain (EC 6.1.1.20)	Phenylalanyl-tRNA synthetase alpha chain (EC 6.1.1.20)
	Phenylalanyl-tRNA synthetase beta chain (EC 6.1.1.20)	Phenylalanyl-tRNA synthetase beta chain (EC 6.1.1.20)
	Phosphate ABC transporter, periplasmic phosphate-binding protein PstS (TC 3.A.1.7.1)	Phosphate ABC transporter, periplasmic phosphate-binding protein PstS (TC 3.A.1.7.1)
	Phosphate acetyltransferase (EC 2.3.1.8)	Phosphate acetyltransferase (EC 2.3.1.8)
	Phosphate regulon sensor protein PhoR (SphS) (EC 2.7.13.3)	Phosphate regulon sensor protein PhoR (SphS) (EC 2.7.13.3)
	Phosphate regulon transcriptional regulatory protein PhoB (SphR)	Phosphate regulon transcriptional regulatory protein PhoB (SphR)
	Phosphate starvation-inducible protein PhoH, predicted ATPase	Phosphate starvation-inducible protein PhoH, predicted ATPase
	Phosphate transport ATP-binding protein PstB (TC 3.A.1.7.1)	Phosphate transport ATP-binding protein PstB (TC 3.A.1.7.1)
	Phosphate transport system permease protein PstA (TC 3.A.1.7.1)	Phosphate transport system permease protein PstA (TC 3.A.1.7.1)
	Phosphate transport system permease protein PstC (TC 3.A.1.7.1)	Phosphate transport system permease protein PstC (TC 3.A.1.7.1)
	Phosphate transport system regulatory protein PhoU	Phosphate transport system regulatory protein PhoU
	Phosphate:acyl-ACP acyltransferase PlsX	Phosphate:acyl-ACP acyltransferase PlsX

Phosphatidate cytidylyltransferase (EC 2.7.7.41)	Phosphatidate cytidylyltransferase (EC 2.7.7.41)
Phosphatidylserine decarboxylase (EC 4.1.1.65)	Phosphatidylserine decarboxylase (EC 4.1.1.65)
Phosphoadenylyl-sulfate reductase [thioredoxin] (EC 1.8.4.8)	Phosphoadenylyl-sulfate reductase [thioredoxin] (EC 1.8.4.8)
Phosphoenolpyruvate carboxykinase [ATP] (EC 4.1.1.49)	Phosphoenolpyruvate carboxykinase [ATP] (EC 4.1.1.49)
Phosphoenolpyruvate carboxylase (EC 4.1.1.31)	Phosphoenolpyruvate carboxylase (EC 4.1.1.31)
Phosphoenolpyruvate carboxylase, archaeal (EC 4.1.1.31)	Phosphoenolpyruvate carboxylase, archaeal (EC 4.1.1.31)
Phosphoenolpyruvate synthase (EC 2.7.9.2)	Phosphoenolpyruvate synthase (EC 2.7.9.2)
Phosphoenolpyruvate-protein phosphotransferase of PTS system (EC 2.7.3.9)	Phosphoenolpyruvate-protein phosphotransferase of PTS system (EC 2.7.3.9)
Phosphoesterase, DHH family protein	Phosphoesterase, DHH family protein
Phosphoglucosamine mutase (EC 5.4.2.10)	Phosphoglucosamine mutase (EC 5.4.2.10)
Phosphoglycerate kinase (EC 2.7.2.3)	Phosphoglycerate kinase (EC 2.7.2.3)
Phospholipid-lipopolysaccharide ABC transporter	Phospholipid-lipopolysaccharide ABC transporter
Phosphomannomutase (EC 5.4.2.8)	Phosphomannomutase (EC 5.4.2.8)
Phosphonate ABC transporter permease protein phnE (TC 3.A.1.9.1)	Phosphonate ABC transporter permease protein phnE (TC 3.A.1.9.1)
Phosphopantetheine adenylyltransferase (EC 2.7.7.3)	Phosphopantetheine adenylyltransferase (EC 2.7.7.3)
Phosphopantetheine adenylyltransferase, type II eukaryotic (EC 2.7.7.3)	Phosphopantetheine adenylyltransferase, type II eukaryotic (EC 2.7.7.3)
Phosphopantothenate synthetase, archaeal	Phosphopantothenate synthetase, archaeal
Phosphopantothenoylcysteine decarboxylase (EC 4.1.1.36)	Phosphopantothenoylcysteine decarboxylase (EC 4.1.1.36)
Phosphopantothenoylcysteine synthetase (EC 6.3.2.5)	Phosphopantothenoylcysteine synthetase (EC 6.3.2.5)
Phosphoribosyl-AMP cyclohydrolase (EC 3.5.4.19)	Phosphoribosyl-AMP cyclohydrolase (EC 3.5.4.19)
Phosphoribosyl-ATP pyrophosphatase (EC 3.6.1.31)	Phosphoribosyl-ATP pyrophosphatase (EC 3.6.1.31)
Phosphoribosylamine--glycine ligase (EC 6.3.4.13)	Phosphoribosylamine--glycine ligase (EC 6.3.4.13)
Phosphoribosylaminoimidazole carboxylase ATPase subunit (EC 4.1.1.21)	Phosphoribosylaminoimidazole carboxylase ATPase subunit (EC 4.1.1.21)
Phosphoribosylaminoimidazole carboxylase catalytic subunit (EC 4.1.1.21)	Phosphoribosylaminoimidazole carboxylase catalytic subunit (EC 4.1.1.21)
Phosphoribosylaminoimidazole-succinocarboxamide	Phosphoribosylaminoimidazole-succinocarboxamide synthase (EC

synthase (EC 6.3.2.6)	6.3.2.6)
Phosphoribosylanthranilate isomerase (EC 5.3.1.24)	Phosphoribosylanthranilate isomerase (EC 5.3.1.24)
Phosphoribosylformimino-5-aminoimidazole carboxamide ribotide isomerase (EC 5.3.1.16)	Phosphoribosylformimino-5-aminoimidazole carboxamide ribotide isomerase (EC 5.3.1.16)
Phosphoribosylformylglycinamidine cyclo-ligase (EC 6.3.3.1)	Phosphoribosylformylglycinamidine cyclo-ligase (EC 6.3.3.1)
Phosphoribosylformylglycinamidine synthase, glutamine amidotransferase subunit (EC 6.3.5.3)	Phosphoribosylformylglycinamidine synthase, glutamine amidotransferase subunit (EC 6.3.5.3)
Phosphoribosylformylglycinamidine synthase, PurS subunit (EC 6.3.5.3)	Phosphoribosylformylglycinamidine synthase, PurS subunit (EC 6.3.5.3)
Phosphoribosylformylglycinamidine synthase, synthetase subunit (EC 6.3.5.3)	Phosphoribosylformylglycinamidine synthase, synthetase subunit (EC 6.3.5.3)
Phosphoribosylglycinamide formyltransferase (EC 2.1.2.2)	Phosphoribosylglycinamide formyltransferase (EC 2.1.2.2)
Phosphoribulokinase (EC 2.7.1.19)	Phosphoribulokinase (EC 2.7.1.19)
Phosphoserine aminotransferase (EC 2.6.1.52)	Phosphoserine aminotransferase (EC 2.6.1.52)
Phosphoserine phosphatase (EC 3.1.3.3)	Phosphoserine phosphatase (EC 3.1.3.3)
Phosphosulfolactate synthase (EC 4.4.1.19)	Phosphosulfolactate synthase (EC 4.4.1.19)
photosystem I P700 chlorophyll a apoprotein subunit Ia (PsaA)	photosystem I P700 chlorophyll a apoprotein subunit Ia (PsaA)
photosystem II protein D1 (PsbA)	photosystem II protein D1 (PsbA)
Phycobilisome core-membrane linker polypeptide	Phycobilisome core-membrane linker polypeptide
Phycobilisome small core linker polypeptide	Phycobilisome small core linker polypeptide
Phytoene dehydrogenase (EC 1.14.99.-)	Phytoene dehydrogenase (EC 1.14.99.-)
Phytoene dehydrogenase and related proteins	Phytoene dehydrogenase and related proteins
Phytoene desaturase (EC 1.14.99.-)	Phytoene desaturase (EC 1.14.99.-)
Phytoene synthase (EC 2.5.1.32)	Phytoene synthase (EC 2.5.1.32)
Polyphosphate kinase (EC 2.7.4.1)	Polyphosphate kinase (EC 2.7.4.1)
Porphobilinogen deaminase (EC 2.5.1.61)	Porphobilinogen deaminase (EC 2.5.1.61)
Possible RuBisCo chaperonin RbcX	Possible RuBisCo chaperonin RbcX
Potassium channel protein	Potassium channel protein
Potassium efflux system KefA protein	Potassium efflux system KefA protein
Potassium uptake protein TrkH	Potassium uptake protein TrkH
Potassium uptake protein, integral membrane component,	Potassium uptake protein, integral membrane component, KtrB

KtrB	
Potassium voltage-gated channel subfamily KQT	Potassium voltage-gated channel subfamily KQT
Predicted ATPase related to phosphate starvation-inducible protein PhoH	Predicted ATPase related to phosphate starvation-inducible protein PhoH
Predicted L-lactate dehydrogenase, Fe-S oxidoreductase subunit YkgE	Predicted L-lactate dehydrogenase, Fe-S oxidoreductase subunit YkgE
Predicted N6-adenine-specific RNA methylase containing THUMP domain	Predicted N6-adenine-specific RNA methylase containing THUMP domain
Predicted RNA-binding protein COG1491	Predicted RNA-binding protein COG1491
Predicted signal transduction protein	Predicted signal transduction protein
Prephenate and/or arogenate dehydrogenase (unknown specificity) (EC 1.3.1.12)(EC 1.3.1.43)	Prephenate and/or arogenate dehydrogenase (unknown specificity) (EC 1.3.1.12)(EC 1.3.1.43)
Prephenate dehydratase (EC 4.2.1.51)	Prephenate dehydratase (EC 4.2.1.51)
Prephenate dehydrogenase (EC 1.3.1.12)	Prephenate dehydrogenase (EC 1.3.1.12)
Probable low-affinity inorganic phosphate transporter	Probable low-affinity inorganic phosphate transporter
Programmed cell death toxin YdcE	Programmed cell death toxin YdcE
Proline dehydrogenase (EC 1.5.99.8) (Proline oxidase)	Proline dehydrogenase (EC 1.5.99.8) (Proline oxidase)
Proline iminopeptidase (EC 3.4.11.5)	Proline iminopeptidase (EC 3.4.11.5)
Proline racemase (EC 5.1.1.4)	Proline racemase (EC 5.1.1.4)
Proline/sodium symporter PutP (TC 2.A.21.2.1)	Proline/sodium symporter PutP (TC 2.A.21.2.1)
Propionyl-CoA carboxylase beta chain (EC 6.4.1.3)	Propionyl-CoA carboxylase beta chain (EC 6.4.1.3)
Proteasome subunit alpha (EC 3.4.25.1), archaeal	Proteasome subunit alpha (EC 3.4.25.1), archaeal
Proteasome subunit beta (EC 3.4.25.1), archaeal	Proteasome subunit beta (EC 3.4.25.1), archaeal
Proteasome-activating AAA-ATPase (PAN), archaeal	Proteasome-activating AAA-ATPase (PAN), archaeal
Protein GerPC, required for proper assembly of spore coat, mutations lead to super-dormant spore	Protein GerPC, required for proper assembly of spore coat, mutations lead to super-dormant spore
Protein RtcB	Protein RtcB
Protein-L-isoaspartate O-methyltransferase (EC 2.1.1.77)	Protein-L-isoaspartate O-methyltransferase (EC 2.1.1.77)
Pterin-4-alpha-carbinolamine dehydratase (EC 4.2.1.96)	Pterin-4-alpha-carbinolamine dehydratase (EC 4.2.1.96)
PTS system, mannitol-specific IIB component (EC 2.7.1.69)	PTS system, mannitol-specific IIB component (EC 2.7.1.69)
PTS system, N-acetylglucosamine-specific IIB component	PTS system, N-acetylglucosamine-specific IIB component (EC

(EC 2.7.1.69)	2.7.1.69)
Purine nucleoside phosphorylase (EC 2.4.2.1)	Purine nucleoside phosphorylase (EC 2.4.2.1)
PurR: transcription regulator associated with purine metabolism	PurR: transcription regulator associated with purine metabolism
Putative 2Fe-2S ferredoxin CbiW involved in B12 biosynthesis	Putative 2Fe-2S ferredoxin CbiW involved in B12 biosynthesis
Putative deoxyribonuclease YcfH	Putative deoxyribonuclease YcfH
putative esterase	putative esterase
Putative formate dehydrogenase oxidoreductase protein	Putative formate dehydrogenase oxidoreductase protein
Putative iron-sulfur cluster assembly scaffold protein for SUF system, SufE2	Putative iron-sulfur cluster assembly scaffold protein for SUF system, SufE2
Putative metal chaperone, involved in Zn homeostasis, GTPase of COG0523 family	Putative metal chaperone, involved in Zn homeostasis, GTPase of COG0523 family
Putative oxidoreductase YncB	Putative oxidoreductase YncB
Putative periplasmic phosphate-binding protein PstS (Halobacteriales type)	Putative periplasmic phosphate-binding protein PstS (Halobacteriales type)
Putative pheromone cAM373 precursor lipoprotein CamS	Putative pheromone cAM373 precursor lipoprotein CamS
Putative preQ0 transporter	Putative preQ0 transporter
putative sodium-dependent bicarbonate transporter	putative sodium-dependent bicarbonate transporter
Putative stomatin/prohibitin-family membrane protease subunit aq_911	Putative stomatin/prohibitin-family membrane protease subunit aq_911
Putative Zn-dependent hydrolase in polyisoprenoid biosynthetic cluster	Putative Zn-dependent hydrolase in polyisoprenoid biosynthetic cluster
Putrescine transport ATP-binding protein PotA (TC 3.A.1.11.1)	Putrescine transport ATP-binding protein PotA (TC 3.A.1.11.1)
Pyridoxamine 5'-phosphate oxidase (EC 1.4.3.5)	Pyridoxamine 5'-phosphate oxidase (EC 1.4.3.5)
Pyridoxine 5'-phosphate synthase (EC 2.6.99.2)	Pyridoxine 5'-phosphate synthase (EC 2.6.99.2)
Pyridoxine biosynthesis glutamine amidotransferase, glutaminase subunit (EC 2.4.2.-)	Pyridoxine biosynthesis glutamine amidotransferase, glutaminase subunit (EC 2.4.2.-)
Pyridoxine biosynthesis glutamine amidotransferase, synthase subunit (EC 2.4.2.-)	Pyridoxine biosynthesis glutamine amidotransferase, synthase subunit (EC 2.4.2.-)
Pyrimidine-nucleoside phosphorylase (EC 2.4.2.2)	Pyrimidine-nucleoside phosphorylase (EC 2.4.2.2)
Pyrophosphate-energized proton pump (EC 3.6.1.1)	Pyrophosphate-energized proton pump (EC 3.6.1.1)
Pyrroline-5-carboxylate reductase (EC 1.5.1.2)	Pyrroline-5-carboxylate reductase (EC 1.5.1.2)
Pyruvate dehydrogenase E1 component alpha subunit (EC	Pyruvate dehydrogenase E1 component alpha subunit (EC 1.2.4.1)

1.2.4.1)	
Pyruvate dehydrogenase E1 component beta subunit (EC 1.2.4.1)	Pyruvate dehydrogenase E1 component beta subunit (EC 1.2.4.1)
Pyruvate kinase (EC 2.7.1.40)	Pyruvate kinase (EC 2.7.1.40)
Pyruvate-flavodoxin oxidoreductase (EC 1.2.7.-)	Pyruvate-flavodoxin oxidoreductase (EC 1.2.7.-)
Pyruvate,phosphate dikinase (EC 2.7.9.1)	Pyruvate,phosphate dikinase (EC 2.7.9.1)
Pyruvate:ferredoxin oxidoreductase, beta subunit (EC 1.2.7.1)	Pyruvate:ferredoxin oxidoreductase, beta subunit (EC 1.2.7.1)
Queuosine Biosynthesis QueC ATPase	Queuosine Biosynthesis QueC ATPase
Queuosine biosynthesis QueD, PTPS-I	Queuosine biosynthesis QueD, PTPS-I
Queuosine Biosynthesis QueE Radical SAM	Queuosine Biosynthesis QueE Radical SAM
Quinolinate phosphoribosyltransferase [decarboxylating] (EC 2.4.2.19)	Quinolinate phosphoribosyltransferase [decarboxylating] (EC 2.4.2.19)
Quinolinate synthetase (EC 2.5.1.72)	Quinolinate synthetase (EC 2.5.1.72)
Radical SAM domain heme biosynthesis protein	Radical SAM domain heme biosynthesis protein
RecA protein	RecA protein
RecD-like DNA helicase YrrC	RecD-like DNA helicase YrrC
Recombination inhibitory protein MutS2	Recombination inhibitory protein MutS2
Replication factor A (ssDNA-binding protein)	Replication factor A (ssDNA-binding protein)
Replication factor C large subunit	Replication factor C large subunit
Replication factor C small subunit	Replication factor C small subunit
Replicative DNA helicase (DnaB) (EC 3.6.4.12)	Replicative DNA helicase (DnaB) (EC 3.6.4.12)
Replicative DNA helicase (EC 3.6.1.-) [SA14-24]	Replicative DNA helicase (EC 3.6.1.-) [SA14-24]
Respiratory nitrate reductase alpha chain (EC 1.7.99.4)	Respiratory nitrate reductase alpha chain (EC 1.7.99.4)
Respiratory nitrate reductase beta chain (EC 1.7.99.4)	Respiratory nitrate reductase beta chain (EC 1.7.99.4)
Riboflavin kinase (EC 2.7.1.26)	Riboflavin kinase (EC 2.7.1.26)
Riboflavin synthase eubacterial/eukaryotic (EC 2.5.1.9)	Riboflavin synthase eubacterial/eukaryotic (EC 2.5.1.9)
Ribokinase (EC 2.7.1.15)	Ribokinase (EC 2.7.1.15)
Ribonuclease BN (EC 3.1.-.-)	Ribonuclease BN (EC 3.1.-.-)
Ribonuclease HII (EC 3.1.26.4)	Ribonuclease HII (EC 3.1.26.4)
Ribonuclease III (EC 3.1.26.3)	Ribonuclease III (EC 3.1.26.3)
Ribonuclease M5 (EC 3.1.26.8)	Ribonuclease M5 (EC 3.1.26.8)

Ribonuclease PH (EC 2.7.7.56)	Ribonuclease PH (EC 2.7.7.56)
Ribonuclease Z (EC 3.1.26.11)	Ribonuclease Z (EC 3.1.26.11)
Ribonucleotide reductase of class Ia (aerobic), alpha subunit (EC 1.17.4.1)	Ribonucleotide reductase of class Ia (aerobic), alpha subunit (EC 1.17.4.1)
Ribonucleotide reductase of class Ia (aerobic), beta subunit (EC 1.17.4.1)	Ribonucleotide reductase of class Ia (aerobic), beta subunit (EC 1.17.4.1)
Ribonucleotide reductase of class II (coenzyme B12-dependent) (EC 1.17.4.1)	Ribonucleotide reductase of class II (coenzyme B12-dependent) (EC 1.17.4.1)
Ribose 5-phosphate isomerase A (EC 5.3.1.6)	Ribose 5-phosphate isomerase A (EC 5.3.1.6)
Ribose ABC transport system, permease protein RbsC (TC 3.A.1.2.1)	Ribose ABC transport system, permease protein RbsC (TC 3.A.1.2.1)
Ribose-1,5-bisphosphate isomerase	Ribose-1,5-bisphosphate isomerase
Ribose-phosphate pyrophosphokinase (EC 2.7.6.1)	Ribose-phosphate pyrophosphokinase (EC 2.7.6.1)
Ribosomal large subunit pseudouridine synthase B (EC 4.2.1.70)	Ribosomal large subunit pseudouridine synthase B (EC 4.2.1.70)
Ribosomal large subunit pseudouridine synthase D (EC 4.2.1.70)	Ribosomal large subunit pseudouridine synthase D (EC 4.2.1.70)
Ribosomal protein L11 methyltransferase (EC 2.1.1.-)	Ribosomal protein L11 methyltransferase (EC 2.1.1.-)
Ribosomal protein S6 glutaminyl transferase	Ribosomal protein S6 glutaminyl transferase
Ribosomal-protein-L7p-serine acetyltransferase	Ribosomal-protein-L7p-serine acetyltransferase
Ribosomal-protein-S5p-alanine acetyltransferase	Ribosomal-protein-S5p-alanine acetyltransferase
Ribosome small subunit-stimulated GTPase EngC	Ribosome small subunit-stimulated GTPase EngC
Ribulose bisphosphate carboxylase large chain (EC 4.1.1.39)	Ribulose bisphosphate carboxylase large chain (EC 4.1.1.39)
Ribulose-phosphate 3-epimerase (EC 5.1.3.1)	Ribulose-phosphate 3-epimerase (EC 5.1.3.1)
RNA polymerase heat shock sigma factor SigI	RNA polymerase heat shock sigma factor SigI
RNA polymerase sigma factor RpoD	RNA polymerase sigma factor RpoD
RNA polymerase sigma factor SigB	RNA polymerase sigma factor SigB
RNA polymerase sigma factor SigW	RNA polymerase sigma factor SigW
RNA polymerase sigma-54 factor RpoN	RNA polymerase sigma-54 factor RpoN
RNA polymerase sporulation specific sigma factor SigF	RNA polymerase sporulation specific sigma factor SigF
RNA polymerase sporulation specific sigma factor SigG	RNA polymerase sporulation specific sigma factor SigG

RNA-binding protein Jag	RNA-binding protein Jag
Rod shape-determining protein MreB	Rod shape-determining protein MreB
Rod shape-determining protein MreC	Rod shape-determining protein MreC
Rod shape-determining protein RodA	Rod shape-determining protein RodA
Rossmann fold nucleotide-binding protein Smf possibly involved in DNA uptake	Rossmann fold nucleotide-binding protein Smf possibly involved in DNA uptake
Rrf2 family transcriptional regulator, group III	Rrf2 family transcriptional regulator, group III
rRNA biogenesis protein Nop5/Nop56	rRNA biogenesis protein Nop5/Nop56
rRNA small subunit methyltransferase I	rRNA small subunit methyltransferase I
RsbS, negative regulator of sigma-B	RsbS, negative regulator of sigma-B
S-adenosylmethionine decarboxylase proenzyme (EC 4.1.1.50), prokaryotic class 1A	S-adenosylmethionine decarboxylase proenzyme (EC 4.1.1.50), prokaryotic class 1A
S-adenosylmethionine synthetase (EC 2.5.1.6)	S-adenosylmethionine synthetase (EC 2.5.1.6)
S-adenosylmethionine:tRNA ribosyltransferase-isomerase (EC 5.-.-.-)	S-adenosylmethionine:tRNA ribosyltransferase-isomerase (EC 5.-.-.-)
Segregation and condensation protein A	Segregation and condensation protein A
Segregation and condensation protein B	Segregation and condensation protein B
Selenocysteine-specific translation elongation factor	Selenocysteine-specific translation elongation factor
Septation ring formation regulator EzrA	Septation ring formation regulator EzrA
Septum site-determining protein MinD	Septum site-determining protein MinD
Serine acetyltransferase (EC 2.3.1.30)	Serine acetyltransferase (EC 2.3.1.30)
Serine hydroxymethyltransferase (EC 2.1.2.1)	Serine hydroxymethyltransferase (EC 2.1.2.1)
Serine phosphatase RsbU, regulator of sigma subunit	Serine phosphatase RsbU, regulator of sigma subunit
Seryl-tRNA synthetase (EC 6.1.1.11)	Seryl-tRNA synthetase (EC 6.1.1.11)
Short chain fatty acids transporter	Short chain fatty acids transporter
Signal peptidase I (EC 3.4.21.89)	Signal peptidase I (EC 3.4.21.89)
Signal peptidase-like protein	Signal peptidase-like protein
Signal peptidase, type IV - prepilin/preflagellin	Signal peptidase, type IV - prepilin/preflagellin
Signal recognition particle receptor protein FtsY (=alpha subunit) (TC 3.A.5.1.1)	Signal recognition particle receptor protein FtsY (=alpha subunit) (TC 3.A.5.1.1)
Signal recognition particle, subunit Ffh SRP54 (TC 3.A.5.1.1)	Signal recognition particle, subunit Ffh SRP54 (TC 3.A.5.1.1)

Signal transduction histidine kinase CheA (EC 2.7.3.-)	Signal transduction histidine kinase CheA (EC 2.7.3.-)
Similar to Hydroxyacylglutathione hydrolase, but in an organism lacking glutathione biosynthesis	Similar to Hydroxyacylglutathione hydrolase, but in an organism lacking glutathione biosynthesis
Sodium-dependent phosphate transporter	Sodium-dependent phosphate transporter
soluble [2Fe-2S] ferredoxin	soluble [2Fe-2S] ferredoxin
SOS-response repressor and protease LexA (EC 3.4.21.88)	SOS-response repressor and protease LexA (EC 3.4.21.88)
Spermidine Putrescine ABC transporter permease component PotB (TC 3.A.1.11.1)	Spermidine Putrescine ABC transporter permease component PotB (TC 3.A.1.11.1)
Spermidine Putrescine ABC transporter permease component potC (TC_3.A.1.11.1)	Spermidine Putrescine ABC transporter permease component potC (TC_3.A.1.11.1)
Spermidine synthase (EC 2.5.1.16)	Spermidine synthase (EC 2.5.1.16)
Spore cortex-lytic enzyme CwlJ	Spore cortex-lytic enzyme CwlJ
Spore maturation protein A	Spore maturation protein A
Spore maturation protein B	Spore maturation protein B
Spore photoproduct lyase (EC 4.1.99.14)	Spore photoproduct lyase (EC 4.1.99.14)
SSU ribosomal protein S10p (S20e)	SSU ribosomal protein S10p (S20e)
SSU ribosomal protein S11e (S17p)	SSU ribosomal protein S11e (S17p)
SSU ribosomal protein S11p (S14e)	SSU ribosomal protein S11p (S14e)
SSU ribosomal protein S13e (S15p)	SSU ribosomal protein S13e (S15p)
SSU ribosomal protein S13p (S18e)	SSU ribosomal protein S13p (S18e)
SSU ribosomal protein S14e (S11p)	SSU ribosomal protein S14e (S11p)
SSU ribosomal protein S15Ae (S8p)	SSU ribosomal protein S15Ae (S8p)
SSU ribosomal protein S15e (S19p)	SSU ribosomal protein S15e (S19p)
SSU ribosomal protein S15p (S13e)	SSU ribosomal protein S15p (S13e)
SSU ribosomal protein S16p	SSU ribosomal protein S16p
SSU ribosomal protein S17p (S11e)	SSU ribosomal protein S17p (S11e)
SSU ribosomal protein S19e	SSU ribosomal protein S19e
SSU ribosomal protein S19p (S15e)	SSU ribosomal protein S19p (S15e)
SSU ribosomal protein S1p	SSU ribosomal protein S1p
SSU ribosomal protein S20e (S10p)	SSU ribosomal protein S20e (S10p)
SSU ribosomal protein S23e (S12p)	SSU ribosomal protein S23e (S12p)

SSU ribosomal protein S24e	SSU ribosomal protein S24e
SSU ribosomal protein S2e (S5p)	SSU ribosomal protein S2e (S5p)
SSU ribosomal protein S3e (S3p)	SSU ribosomal protein S3e (S3p)
SSU ribosomal protein S3p (S3e)	SSU ribosomal protein S3p (S3e)
SSU ribosomal protein S4p (S9e)	SSU ribosomal protein S4p (S9e)
SSU ribosomal protein S5p (S2e)	SSU ribosomal protein S5p (S2e)
SSU ribosomal protein S7p (S5e)	SSU ribosomal protein S7p (S5e)
SSU ribosomal protein S8p (S15Ae)	SSU ribosomal protein S8p (S15Ae)
SSU ribosomal protein S9e (S4p)	SSU ribosomal protein S9e (S4p)
SSU ribosomal protein S9p (S16e)	SSU ribosomal protein S9p (S16e)
SSU ribosomal protein SAe (S2p)	SSU ribosomal protein SAe (S2p)
Stage 0 sporulation two-component response regulator (Spo0A)	Stage 0 sporulation two-component response regulator (Spo0A)
Stage II sporulation protein B	Stage II sporulation protein B
Stage II sporulation serine phosphatase for sigma-F activation (SpoIIE)	Stage II sporulation serine phosphatase for sigma-F activation (SpoIIE)
Stage III sporulation protein AA	Stage III sporulation protein AA
Stage III sporulation protein AB	Stage III sporulation protein AB
Stage III sporulation protein AF	Stage III sporulation protein AF
Stage III sporulation protein AG	Stage III sporulation protein AG
Stage IV sporulation pro-sigma-K processing enzyme (SpoIVFB)	Stage IV sporulation pro-sigma-K processing enzyme (SpoIVFB)
Stage IV sporulation protein A	Stage IV sporulation protein A
Stage IV sporulation protein B	Stage IV sporulation protein B
Stage IV sporulation protein FA (SpoIVFA)	Stage IV sporulation protein FA (SpoIVFA)
Stage V sporulation protein AA (SpoVAA)	Stage V sporulation protein AA (SpoVAA)
Stage V sporulation protein AE (SpoVAE)	Stage V sporulation protein AE (SpoVAE)
Stage V sporulation protein AF (SpoVAF)	Stage V sporulation protein AF (SpoVAF)
Stage V sporulation protein B	Stage V sporulation protein B
Stage V sporulation protein involved in spore cortex synthesis (SpoVR)	Stage V sporulation protein involved in spore cortex synthesis (SpoVR)
Stage V sporulation protein whose disruption leads to the production of immature spores (SpoVK)	Stage V sporulation protein whose disruption leads to the production of immature spores (SpoVK)
Substrate-specific component BioY of biotin ECF transporter	Substrate-specific component BioY of biotin ECF transporter

Substrate-specific component RibU of riboflavin ECF transporter	Substrate-specific component RibU of riboflavin ECF transporter
Substrate-specific component ThiT of thiamin ECF transporter	Substrate-specific component ThiT of thiamin ECF transporter
Succinate dehydrogenase cytochrome b subunit	Succinate dehydrogenase cytochrome b subunit
Succinate dehydrogenase cytochrome b-556 subunit	Succinate dehydrogenase cytochrome b-556 subunit
Succinate dehydrogenase cytochrome b558 subunit	Succinate dehydrogenase cytochrome b558 subunit
Succinate dehydrogenase flavoprotein subunit (EC 1.3.99.1)	Succinate dehydrogenase flavoprotein subunit (EC 1.3.99.1)
Succinate dehydrogenase iron-sulfur protein (EC 1.3.99.1)	Succinate dehydrogenase iron-sulfur protein (EC 1.3.99.1)
Succinyl-CoA ligase [ADP-forming] alpha chain (EC 6.2.1.5)	Succinyl-CoA ligase [ADP-forming] alpha chain (EC 6.2.1.5)
Succinyl-CoA ligase [ADP-forming] beta chain (EC 6.2.1.5)	Succinyl-CoA ligase [ADP-forming] beta chain (EC 6.2.1.5)
Succinyl-CoA:3-ketoacid-coenzyme A transferase subunit A (EC 2.8.3.5)	Succinyl-CoA:3-ketoacid-coenzyme A transferase subunit A (EC 2.8.3.5)
Succinyl-CoA:3-ketoacid-coenzyme A transferase subunit B (EC 2.8.3.5)	Succinyl-CoA:3-ketoacid-coenzyme A transferase subunit B (EC 2.8.3.5)
Sucrose-6-phosphate hydrolase (EC 3.2.1.B3)	Sucrose-6-phosphate hydrolase (EC 3.2.1.B3)
Sulfate adenylyltransferase subunit 2 (EC 2.7.7.4)	Sulfate adenylyltransferase subunit 2 (EC 2.7.7.4)
Sulfate and thiosulfate import ATP-binding protein CysA (EC 3.6.3.25)	**Sulfate and thiosulfate import ATP-binding protein CysA (EC 3.6.3.25)**
Sulfate permease	Sulfate permease
Sulfate permease, Trk-type	Sulfate permease, Trk-type
Sulfite reductase [NADPH] flavoprotein alpha-component (EC 1.8.1.2)	Sulfite reductase [NADPH] flavoprotein alpha-component (EC 1.8.1.2)
Sulfite reductase [NADPH] hemoprotein beta-component (EC 1.8.1.2)	Sulfite reductase [NADPH] hemoprotein beta-component (EC 1.8.1.2)
Sulfur carrier protein adenylyltransferase ThiF	Sulfur carrier protein adenylyltransferase ThiF
Superoxide dismutase [Mn] (EC 1.15.1.1)	Superoxide dismutase [Mn] (EC 1.15.1.1)
Survival protein SurA precursor (Peptidyl-prolyl cis-trans isomerase SurA) (EC 5.2.1.8)	Survival protein SurA precursor (Peptidyl-prolyl cis-trans isomerase SurA) (EC 5.2.1.8)
SusD, outer membrane protein	SusD, outer membrane protein
TATA-box binding protein	TATA-box binding protein
Thermosome subunit	Thermosome subunit
Thiamin ABC transporter, transmembrane component	Thiamin ABC transporter, transmembrane component

Thiamin pyrophosphokinase (EC 2.7.6.2)	Thiamin pyrophosphokinase (EC 2.7.6.2)
Thiamin-phosphate pyrophosphorylase (EC 2.5.1.3)	Thiamin-phosphate pyrophosphorylase (EC 2.5.1.3)
Thiamin-phosphate synthase ThiN (EC 2.5.1.3)	Thiamin-phosphate synthase ThiN (EC 2.5.1.3)
Thiaminase II (EC 3.5.99.2) involved in salvage of thiamin pyrimidine moiety	Thiaminase II (EC 3.5.99.2) involved in salvage of thiamin pyrimidine moiety
Thiamine biosynthesis protein thiI	Thiamine biosynthesis protein thiI
Thiamine-monophosphate kinase (EC 2.7.4.16)	Thiamine-monophosphate kinase (EC 2.7.4.16)
Thiazole biosynthesis protein ThiG	Thiazole biosynthesis protein ThiG
Thiol peroxidase, Tpx-type (EC 1.11.1.15)	Thiol peroxidase, Tpx-type (EC 1.11.1.15)
Thioredoxin reductase (EC 1.8.1.9)	Thioredoxin reductase (EC 1.8.1.9)
Thiosulfate sulfurtransferase, rhodanese (EC 2.8.1.1)	Thiosulfate sulfurtransferase, rhodanese (EC 2.8.1.1)
Threonine dehydratase (EC 4.3.1.19)	Threonine dehydratase (EC 4.3.1.19)
Threonine dehydratase, catabolic (EC 4.3.1.19)	Threonine dehydratase, catabolic (EC 4.3.1.19)
Threonine dehydrogenase and related Zn-dependent dehydrogenases	Threonine dehydrogenase and related Zn-dependent dehydrogenases
Threonine synthase (EC 4.2.3.1)	Threonine synthase (EC 4.2.3.1)
Threonyl-tRNA synthetase (EC 6.1.1.3)	Threonyl-tRNA synthetase (EC 6.1.1.3)
Thymidine kinase (EC 2.7.1.21)	Thymidine kinase (EC 2.7.1.21)
Thymidylate kinase (EC 2.7.4.9)	Thymidylate kinase (EC 2.7.4.9)
Thymidylate synthase (EC 2.1.1.45)	Thymidylate synthase (EC 2.1.1.45)
Thymidylate synthase thyX (EC 2.1.1.-)	Thymidylate synthase thyX (EC 2.1.1.-)
TldD family protein, Actinobacterial subgroup	TldD family protein, Actinobacterial subgroup
TldE/PmbA family protein, Actinobacterial subgroup	TldE/PmbA family protein, Actinobacterial subgroup
tmRNA-binding protein SmpB	tmRNA-binding protein SmpB
tolB protein precursor, periplasmic protein involved in the tonb-independent uptake of group A colicins	tolB protein precursor, periplasmic protein involved in the tonb-independent uptake of group A colicins
TonB-dependent receptor	TonB-dependent receptor
Topoisomerase IV subunit A (EC 5.99.1.-)	Topoisomerase IV subunit A (EC 5.99.1.-)
Topoisomerase IV subunit B (EC 5.99.1.-)	Topoisomerase IV subunit B (EC 5.99.1.-)

TPR-repeat-containing protein, putative component of Menaquinone-cytochrome C reductase	TPR-repeat-containing protein, putative component of Menaquinone-cytochrome C reductase
Transaldolase (EC 2.2.1.2)	Transaldolase (EC 2.2.1.2)
Transcription antitermination protein NusG	Transcription antitermination protein NusG
Transcription elongation factor GreA	Transcription elongation factor GreA
Transcription factor S	Transcription factor S
Transcription initiation factor B	Transcription initiation factor B
Transcription regulator Bat, regulates bacteriorhodopsin synthesis	Transcription regulator Bat, regulates bacteriorhodopsin synthesis
Transcription termination factor Rho	Transcription termination factor Rho
Transcription termination protein NusA	Transcription termination protein NusA
Transcription termination protein NusB	Transcription termination protein NusB
Transcription-repair coupling factor	Transcription-repair coupling factor
Transcriptional regulator, MerR family	Transcriptional regulator, MerR family
Transcriptional repressor for NAD biosynthesis in gram-positives	Transcriptional repressor for NAD biosynthesis in gram-positives
Transketolase, N-terminal section (EC 2.2.1.1)	Transketolase, N-terminal section (EC 2.2.1.1)
Translation elongation factor 1 alpha subunit	Translation elongation factor 1 alpha subunit
Translation elongation factor 1 beta subunit	Translation elongation factor 1 beta subunit
Translation elongation factor 2	Translation elongation factor 2
Translation elongation factor G	Translation elongation factor G
Translation elongation factor G-related protein	Translation elongation factor G-related protein
Translation elongation factor LepA	Translation elongation factor LepA
Translation elongation factor P	Translation elongation factor P
Translation elongation factor Tu	Translation elongation factor Tu
Translation initiation factor 1	Translation initiation factor 1
Translation initiation factor 1A	Translation initiation factor 1A
Translation initiation factor 2	Translation initiation factor 2
Translation initiation factor 2B alpha/beta/delta-type subunit	Translation initiation factor 2B alpha/beta/delta-type subunit
Translation initiation factor SUI1-related protein	Translation initiation factor SUI1-related protein
Transmembrane component of general energizing module of ECF transporters	Transmembrane component of general energizing module of ECF transporters

Trehalose synthase (EC 5.4.99.16)	Trehalose synthase (EC 5.4.99.16)
Trehalose-6-phosphate hydrolase (EC 3.2.1.93)	Trehalose-6-phosphate hydrolase (EC 3.2.1.93)
Tricarboxylate transport membrane protein TctA	Tricarboxylate transport membrane protein TctA
Triosephosphate isomerase (EC 5.3.1.1)	Triosephosphate isomerase (EC 5.3.1.1)
Trk system potassium uptake protein TrkA	Trk system potassium uptake protein TrkA
TrkA-N:Sodium/hydrogen exchanger	TrkA-N:Sodium/hydrogen exchanger
tRNA (5-methylaminomethyl-2-thiouridylate)-methyltransferase (EC 2.1.1.61)	tRNA (5-methylaminomethyl-2-thiouridylate)-methyltransferase (EC 2.1.1.61)
tRNA (Guanine37-N1) -methyltransferase (EC 2.1.1.31)	tRNA (Guanine37-N1) -methyltransferase (EC 2.1.1.31)
tRNA nucleotidyltransferase (EC 2.7.7.21) (EC 2.7.7.25)	tRNA nucleotidyltransferase (EC 2.7.7.21) (EC 2.7.7.25)
tRNA nucleotidyltransferase, archaeal type (EC 2.7.7.21) (EC 2.7.7.25)	tRNA nucleotidyltransferase, archaeal type (EC 2.7.7.21) (EC 2.7.7.25)
tRNA pseudouridine 13 synthase (EC 4.2.1.-)	tRNA pseudouridine 13 synthase (EC 4.2.1.-)
tRNA pseudouridine 55 synthase (EC 4.2.1.70)	tRNA pseudouridine 55 synthase (EC 4.2.1.70)
tRNA pseudouridine synthase A (EC 4.2.1.70)	tRNA pseudouridine synthase A (EC 4.2.1.70)
tRNA-guanine transglycosylase (EC 2.4.2.29)	tRNA-guanine transglycosylase (EC 2.4.2.29)
tRNA-i(6)A37 methylthiotransferase	tRNA-i(6)A37 methylthiotransferase
tRNA-intron endonuclease (EC 3.1.27.9)	tRNA-intron endonuclease (EC 3.1.27.9)
tRNA:Cm32/Um32 methyltransferase	tRNA:Cm32/Um32 methyltransferase
tRNA(Ile)-lysidine synthetase (EC 6.3.4.19)	tRNA(Ile)-lysidine synthetase (EC 6.3.4.19)
Tryptophan 2,3-dioxygenase (EC 1.13.11.11)	Tryptophan 2,3-dioxygenase (EC 1.13.11.11)
Tryptophan synthase alpha chain (EC 4.2.1.20)	Tryptophan synthase alpha chain (EC 4.2.1.20)
Tryptophan synthase beta chain (EC 4.2.1.20)	Tryptophan synthase beta chain (EC 4.2.1.20)
Tryptophanyl-tRNA synthetase (EC 6.1.1.2)	Tryptophanyl-tRNA synthetase (EC 6.1.1.2)
TsaB protein, required for threonylcarbamoyladenosine (t(6)A) formation in tRNA	TsaB protein, required for threonylcarbamoyladenosine (t(6)A) formation in tRNA
TsaC protein (YrdC domain) required for threonylcarbamoyladenosine t(6)A37 modification in tRNA	TsaC protein (YrdC domain) required for threonylcarbamoyladenosine t(6)A37 modification in tRNA
TsaD/Kae1/Qri7 protein, required for threonylcarbamoyladenosine t(6)A37 formation in tRNA	TsaD/Kae1/Qri7 protein, required for threonylcarbamoyladenosine t(6)A37 formation in tRNA
Twin-arginine translocation protein TatC	Twin-arginine translocation protein TatC

Two-component sensor histidine kinase, malate (EC 2.7.3.-)	Two-component sensor histidine kinase, malate (EC 2.7.3.-)
Type I restriction-modification system, DNA-methyltransferase subunit M (EC 2.1.1.72)	Type I restriction-modification system, DNA-methyltransferase subunit M (EC 2.1.1.72)
Type I restriction-modification system, restriction subunit R (EC 3.1.21.3)	Type I restriction-modification system, restriction subunit R (EC 3.1.21.3)
Type II secretory pathway, ATPase PulE/Tfp pilus assembly pathway, ATPase PilB	Type II secretory pathway, ATPase PulE/Tfp pilus assembly pathway, ATPase PilB
Type III restriction-modification system methylation subunit (EC 2.1.1.72)	Type III restriction-modification system methylation subunit (EC 2.1.1.72)
Type IV fimbrial assembly protein PilC	Type IV fimbrial assembly protein PilC
Type IV fimbrial assembly, ATPase PilB	Type IV fimbrial assembly, ATPase PilB
Type IV pilin PilA	Type IV pilin PilA
Tyrosyl-tRNA synthetase (EC 6.1.1.1)	Tyrosyl-tRNA synthetase (EC 6.1.1.1)
Ubiquinol--cytochrome c reductase, cytochrome B subunit (EC 1.10.2.2)	Ubiquinol--cytochrome c reductase, cytochrome B subunit (EC 1.10.2.2)
Ubiquinone biosynthesis monooxygenase UbiB	Ubiquinone biosynthesis monooxygenase UbiB
Ubiquinone/menaquinone biosynthesis methyltransferase UbiE (EC 2.1.1.-)	Ubiquinone/menaquinone biosynthesis methyltransferase UbiE (EC 2.1.1.-)
UDP-glucose 4-epimerase (EC 5.1.3.2)	UDP-glucose 4-epimerase (EC 5.1.3.2)
UDP-glucose dehydrogenase (EC 1.1.1.22)	UDP-glucose dehydrogenase (EC 1.1.1.22)
UDP-N-acetylenolpyruvoylglucosamine reductase (EC 1.1.1.158)	UDP-N-acetylenolpyruvoylglucosamine reductase (EC 1.1.1.158)
UDP-N-acetylglucosamine 1-carboxyvinyltransferase (EC 2.5.1.7)	UDP-N-acetylglucosamine 1-carboxyvinyltransferase (EC 2.5.1.7)
UDP-N-acetylglucosamine 2-epimerase (EC 5.1.3.14)	UDP-N-acetylglucosamine 2-epimerase (EC 5.1.3.14)
UDP-N-acetylglucosamine--N-acetylmuramyl-(pentapeptide) pyrophosphoryl-undecaprenol N-acetylglucosamine transferase (EC 2.4.1.227)	UDP-N-acetylglucosamine--N-acetylmuramyl-(pentapeptide) pyrophosphoryl-undecaprenol N-acetylglucosamine transferase (EC 2.4.1.227)
UDP-N-acetylmuramoylalanyl-D-glutamate--2,6-diaminopimelate ligase (EC 6.3.2.13)	UDP-N-acetylmuramoylalanyl-D-glutamate--2,6-diaminopimelate ligase (EC 6.3.2.13)
UDP-N-acetylmuramoylalanyl-D-glutamyl-2,6-diaminopimelate--D-alanyl-D-alanine ligase (EC 6.3.2.10)	UDP-N-acetylmuramoylalanyl-D-glutamyl-2,6-diaminopimelate--D-alanyl-D-alanine ligase (EC 6.3.2.10)
Undecaprenyl-phosphate galactosephosphotransferase (EC 2.7.8.6)	Undecaprenyl-phosphate galactosephosphotransferase (EC 2.7.8.6)
Undecaprenyl-phosphate N-acetylglucosaminyl 1-phosphate	Undecaprenyl-phosphate N-acetylglucosaminyl 1-phosphate

transferase (EC 2.7.8.-)	transferase (EC 2.7.8.-)
Universal stress protein family	Universal stress protein family
UPF0234 protein YajQ	UPF0234 protein YajQ
Uracil DNA glycosylase superfamily protein	Uracil DNA glycosylase superfamily protein
Uracil phosphoribosyltransferase (EC 2.4.2.9)	Uracil phosphoribosyltransferase (EC 2.4.2.9)
Uracil-DNA glycosylase, family 4	Uracil-DNA glycosylase, family 4
Uracil-xanthine permease	Uracil-xanthine permease
Urea ABC transporter, ATPase protein UrtE	Urea ABC transporter, ATPase protein UrtE
Urea ABC transporter, permease protein UrtB	Urea ABC transporter, permease protein UrtB
Urea ABC transporter, urea binding protein	Urea ABC transporter, urea binding protein
Urea carboxylase-related ABC transporter, ATPase protein	Urea carboxylase-related ABC transporter, ATPase protein
Urease accessory protein UreG	Urease accessory protein UreG
Urease alpha subunit (EC 3.5.1.5)	Urease alpha subunit (EC 3.5.1.5)
Urease gamma subunit (EC 3.5.1.5)	Urease gamma subunit (EC 3.5.1.5)
Urocanate hydratase (EC 4.2.1.49)	Urocanate hydratase (EC 4.2.1.49)
Uroporphyrinogen-III methyltransferase (EC 2.1.1.107)	Uroporphyrinogen-III methyltransferase (EC 2.1.1.107)
V-type ATP synthase subunit A (EC 3.6.3.14)	V-type ATP synthase subunit A (EC 3.6.3.14)
V-type ATP synthase subunit B (EC 3.6.3.14)	V-type ATP synthase subunit B (EC 3.6.3.14)
V-type ATP synthase subunit D (EC 3.6.3.14)	V-type ATP synthase subunit D (EC 3.6.3.14)
V-type ATP synthase subunit I (EC 3.6.3.14)	V-type ATP synthase subunit I (EC 3.6.3.14)
Various polyols ABC transporter, permease component 1	Various polyols ABC transporter, permease component 1
Vitamin B12 ABC transporter, ATPase component BtuD	Vitamin B12 ABC transporter, ATPase component BtuD
Vitamin B12 ABC transporter, B12-binding component BtuF	Vitamin B12 ABC transporter, B12-binding component BtuF
Vitamin B12 ABC transporter, permease component BtuC	Vitamin B12 ABC transporter, permease component BtuC
Xanthine-guanine phosphoribosyltransferase (EC 2.4.2.22)	Xanthine-guanine phosphoribosyltransferase (EC 2.4.2.22)
Xanthine/uracil/thiamine/ascorbate permease family protein	Xanthine/uracil/thiamine/ascorbate permease family protein
Xylulose kinase (EC 2.7.1.17)	Xylulose kinase (EC 2.7.1.17)
Xylulose-5-phosphate phosphoketolase (EC 4.1.2.9)	Xylulose-5-phosphate phosphoketolase (EC 4.1.2.9)
YjbH-like, GTP pyrophosphokinase domain	YjbH-like, GTP pyrophosphokinase domain
YpfJ protein, zinc metalloprotease superfamily	YpfJ protein, zinc metalloprotease superfamily

Zinc ABC transporter, inner membrane permease protein ZnuB	Zinc ABC transporter, inner membrane permease protein ZnuB
Zinc metalloproteinase precursor (EC 3.4.24.29)	Zinc metalloproteinase precursor (EC 3.4.24.29)
Zinc uptake regulation protein ZUR	Zinc uptake regulation protein ZUR

Appendix 5: List of subsystems in Carbohydrate metabolism

Subsystems	No. of assigned bases (b)						
	DAL	GAL	BLA	LADP5	LADP6	LA	MUP
Acetyl-CoA fermentation to Butyrate	52600	173000	114000	14100	27500	19700	13100
Ethanolamine utilization	5480	133000	7180	0	1830	0	0
Mannose Metabolism	16600	107000	20800	6020	9510	8750	876
L-rhamnose utilization	8420	107000	15400	222	0	0	237
Acetoin butanediol metabolism	7990	106000	19200	1740	5630	4940	5160
Sugar utilization in Thermotogales	136000	99400	213000	42000	55400	43700	12800
Acetone Butanol Ethanol Synthesis	31500	78100	71400	14200	27500	16800	11200
Lactose utilization	7710	72000	7190	0	0	0	0
Lactose and Galactose Uptake and Utilization	17600	56900	29500	4820	7240	4480	1110
Isobutyryl-CoA to Propionyl-CoA Module	21500	55200	42200	7560	19700	11100	6240
Malonate decarboxylase	1680	51000	4760	0	0	0	0
2-Ketogluconate Utilization	816	49000	933	0	0	0	0
Glycerol fermentation to 1	6670	46600	7720	1380	2390	2510	1030
Trehalose Uptake and Utilization	1900	46600	6370	225	330	0	0
Entner-Doudoroff Pathway	64700	43000	117000	16800	26600	22300	16100
Pentose phosphate pathway	24000	41300	40600	11500	17500	12300	3110
Pentose phosphate pathway in plants	17600	40600	25100	7170	13200	8140	2680
Fermentations in Streptococci	48100	40300	80200	15900	23200	13300	3060
Butanol Biosynthesis	33300	37900	69000	10800	24400	14600	8860
Pyruvate metabolism II: acetyl-CoA, acetogenesis from pyruvate	70000	36500	153000	14900	39400	18600	13300
D-Sorbitol(D-Glucitol) and L-Sorbose Utilization	3280	36000	3800	0	0	0	0
Methanogenesis	0	35700	222	0	459	0	0
millsd methanogensis	0	35700	222	0	459	0	0
Carboxysome	5400	34800	12700	1690	2910	4510	444
Photorespiration (oxidative C2 cycle) in plants	27100	32000	44900	12200	16400	14700	6680
Dehydrogenase complexes	36500	31600	59100	13100	25600	13200	5990
Fermentations: Lactate	25000	29500	44500	6300	11500	3890	1110
Rubisco shunt in plants	16600	29400	28100	8970	14500	13600	4240
Fermentations: Mixed acid	21000	28100	38400	6950	13800	6940	1070
Calvin-Benson cycle	27600	26000	40400	9540	15500	11800	4610
Calvin-Benson-Bassham cycle in plants	16300	25800	23900	4360	8270	7210	3070
Serine-glyoxylate cycle	113000	24800	191000	44200	71700	55700	31000
L-Arabinose utilization	14500	24700	19400	1290	1010	1880	0
L-ascorbate utilization (and related gene clusters)	870	22700	3800	0	0	0	0
Maltose and Maltodextrin Utilization	25600	20500	51600	10900	22800	7540	1680
Methylcitrate cycle	26900	19500	52400	5230	12600	8280	7290
Propionate-CoA to Succinate Module	26200	19500	28900	9960	14800	10700	7530
L-Arabinose CS	14600	18600	21400	843	1200	1880	0
Ethylmalonyl-CoA pathway of C2 assimilation, GJO	11100	18500	19200	342	2760	897	2530
Ethylmalonyl-CoA pathway of C2 assimilation	11100	18500	18900	342	2760	897	2530

CO2 uptake carboxysome	4400	18100	10800	3190	7740	8170	1870
beta carboxysome	4130	17900	9200	465	1370	2400	0
Photorespiration (oxidative C2 cycle)	47300	15600	84600	13400	22400	18900	10900
Methanogenesis from methylated compounds	0	15600	0	0	0	0	0
Alpha-Amylase locus in Streptocococcus	2100	15100	4810	0	720	1210	0
Glycolysis and Gluconeogenesis in Streptococci	55600	15000	88000	16000	23100	16800	7780
Glycolysis and Gluconeogenesis	51400	15000	83500	17800	26400	22700	11900
Chitin and N-acetylglucosamine utilization	21700	14200	18400	2440	5900	4460	0
Acetyl-CoA biosynthesis in plants	19300	13700	42200	10600	14400	6220	5230
alpha carboxysome	9810	12500	20500	11600	20200	17500	4830
Formaldehyde assimilation: Ribulose monophosphate pathway	30500	12400	45600	7150	10800	6750	6800
TCA cycle in plants	66700	11700	119000	32300	46400	35600	16500
Methylglyoxal Metabolism	20300	11000	48600	3560	7510	5990	5230
Propanediol utilization	4610	10300	4700	273	1650	240	219
Xylose utilization	9750	9690	18000	0	876	2530	447
D-ribose utilization	12700	9450	22500	1240	1800	1350	243
Dihydroxyacetone kinases	5020	8460	3810	0	0	0	0
Sucrose metabolism in plants	41700	8440	72400	19300	28700	22800	9580
Tricarballylate Utilization	1640	8270	4840	0	0	0	0
Glycolysis test	46400	7600	74000	16300	24900	22200	11600
Glycolysis and Gluconeogenesis, including Archaeal enzymes	31300	6730	58700	14900	22700	20300	11500
Trehalose Biosynthesis	23700	6720	37000	4810	14800	5540	1280
Fructooligosaccharides(FOS) and Raffinose Utilization	9020	6690	12700	489	5370	1490	345
Glycolysis and Gluconeogenesis in plants	62600	6520	104000	26600	40400	36100	15400
Galactose degradation in plants	9990	6520	23800	4820	7150	3860	1510
Starch degradation in plants	5980	6520	8710	576	1710	531	663
Trehalose metabolism in plants	5680	6520	11000	1460	804	255	663
Hexose Phosphate Uptake System	826	4770	2520	0	0	0	0
Pyruvate Alanine Serine Interconversions	27800	4650	49100	4730	7310	3870	2930
Fructose utilization	11800	4060	12000	2360	5530	3980	918
EC 5.1.3.- Racemases and epimerases acting on carbohydrates and derivatives	8800	3890	21200	3880	6010	3870	846
D-Galacturonate and D-Glucuronate Utilization	20000	3830	33200	3090	4860	1630	1190
D-galactonate catabolism	3150	3770	1670	0	0	0	0
Citrate Metabolism KE	23300	3580	38300	14900	21600	20800	4760
One-carbon metabolism by tetrahydropterines	8680	3480	14700	5920	7100	6590	1980
Propionyl-CoA to Succinyl-CoA Module	8400	3350	19400	6320	6710	6630	1720
Glycogen metabolism	23400	3090	32800	2050	8130	2990	663
Glycerate metabolism	11900	2860	31600	2510	4650	2070	510
Pyruvate metabolism I: anaplerotic reactions, PEP in Mycobacteria	63200	2760	113000	24700	40900	33100	10600
D-gluconate and ketogluconates metabolism	15700	2480	20300	237	1110	1320	0
Alpha-acetolactate operon	225	2370	312	0	0	0	0
Glycerol and Glycerol-3-phosphate Uptake and Utilization, 3-	18700	2190	43000	15100	17800	14500	4990

propanediol							
Pyruvate metabolism I: anaplerotic reactions, PEP	37500	1960	73100	16400	32000	18500	6060
D-galactarate, D-glucarate and D-glycerate catabolism	9070	1930	18100	315	966	408	237
D-galactarate, D-glucarate and D-glycerate catabolism - gjo	9070	1930	18100	315	966	408	237
Inositol catabolism	5350	1910	12100	0	0	0	0
Peripheral Glucose Catabolism Pathways	3300	1090	2810	213	0	570	0
Mannitol Utilization	9460	1050	11600	1080	9940	591	534
Citrate Metabolism KE4	41800	933	60900	14800	22800	22000	8390
acinetobacter tca	87500	798	139000	30700	53500	46800	19300
TCA Cycle	87500	798	139000	30700	53500	46800	19300
Folate-mediated one-carbon metabolism in plants	29800	561	44600	16900	23400	22200	9050
Glycolate, glyoxylate interconversions	9260	498	25700	228	1280	441	0
Glyoxylate bypass	37800	165	59900	15600	22900	15900	9670
Quinones HGM	49400	0	76900	20800	29400	29700	11100
Citrate Metabolism KE3	35500	0	53200	14500	21000	21100	8180
Lactate utilization temp	21700	0	37400	3680	4290	4520	1300
Lactate utilization	15400	0	23600	888	720	678	786
Lacto-N-Biose I and Galacto-N-Biose Metabolic Pathway	12000	0	19800	2490	5070	2670	846
Deoxyribose and Deoxynucleoside Catabolism	9750	0	12300	7000	9190	5230	1140
N-Acetyl-Galactosamine and Galactosamine Utilization	9030	0	7630	2630	6220	5960	567
Starch biosynthesis in plants	8030	0	13600	1100	2160	2090	612
2-O-alpha-mannosyl-D-glycerate utilization	6250	0	6050	0	0	0	0
Beta-Glucoside Metabolism	5930	0	14100	0	0	0	0
D-Tagatose and Galactitol Utilization	3850	0	5370	438	1110	1740	1010
Xylose utilization in plants	2930	0	3470	0	876	2530	213
D-mannitol and D-mannose degradation in plants	2530	0	4670	0	0	0	0
Pyruvate:ferredoxin oxidoreductase	2350	0	3330	1870	4010	1490	798
Melibiose Utilization	1700	0	2480	0	1490	336	0
Unknown carbohydrate utilization (cluster Yeg)	1620	0	2290	744	951	1570	597
Melibiose degradation in plants	1490	0	1580	0	1490	336	0
L-fucose utilization temp	1030	0	1300	0	0	0	0
Citrate Metabolism KE2	985	0	1370	231	0	189	0
Sucrose utilization	822	0	2410	0	2990	222	345
beta-glucuronide utilization	783	0	2140	0	0	0	0
Unknown sugar utilization (cluster yphABCDEFG)	699	0	2370	0	0	0	0
Unknown carbohydrate utilization (cluster Ydj)	522	0	828	0	0	0	0
L-fucose utilization	471	0	4610	0	0	0	0
Sucrose utilization Shewanella	366	0	2410	0	0	0	0
Conserved cluster around inner membrane protein gene yghQ probably involved in polysaccharide biosynthesis	336	0	748	0	0	0	0
Erythritol utilization	231	0	2500	0	0	0	0
Tagatose utilization	222	0	318	0	0	534	264
D-allose utilization	213	0	1500	0	0	0	0
Soluble methane monooxygenase (sMMO)	213	0	722	0	0	0	0

Acetone carboxylase	0	0	0	3470	4490	9400	951
Cellulosome	0	0	0	0	714	0	0
CitAB	0	0	758	0	0	0	0
VC0266	0	0	541	0	0	0	0
Chitobiose	0	0	153	0	0	0	0

Appendix 6: Protein families predicted in subsystems of Nitrogen metabolism

Subsystem: Amidase clustered with urea and nitrile hydratase functions	Number of aligned bases (b)						
	DAL	GAL	BLA	LADP5	LADP6	LA	MUP
Amidase clustered with urea ABC transporter and nitrile hydratase functions	0	0	1.17E+03	0	0	0	0

Subsystem: Ammonia assimilation	Number of aligned bases (b)						
	DAL	GAL	BLA	LADP5	LADP6	LA	MUP
Ammonium transporter	4.11E+02	0.00E+00	8.17E+03	0.00E+00	1.07E+03	9.33E+02	0.00E+00
Ferredoxin-dependent glutamate synthase (EC 1.4.7.1)	1.09E+03	6.13E+03	2.29E+03	0	0	0	0
Glutamate synthase [NADPH] large chain (EC 1.4.1.13)	1.26E+04	1.39E+03	3.00E+04	9.84E+02	7.68E+03	2.59E+03	2.66E+03
Glutamate synthase [NADPH] putative GlxC chain (EC 1.4.1.13)	0.00E+00	2.87E+03	0	0	0	0	0
Glutamate synthase [NADPH] small chain (EC 1.4.1.13)	4.71E+03	0.00E+00	7.50E+03	5.87E+02	0.00E+00	5.70E+02	4.95E+02
Glutamate-ammonia-ligase adenylyltransferase (EC 2.7.7.42)	2.51E+03	2.57E+03	7.04E+03	0	0	0	0
Glutamine amidotransferase, class-II	0.00E+00	2.31E+02	0	0	0	0	0
Glutamine synthetase inactivating factor IF17	0.00E+00	5.66E+03	0	0	0	0	0
Glutamine synthetase type I (EC 6.3.1.2)	7.03E+03	0.00E+00	7.40E+03	4.08E+03	6.75E+03	6.41E+03	3.39E+03
Glutamine synthetase type II, eukaryotic (EC 6.3.1.2)	8.67E+02	1.12E+04	0	0	0	0	0
Glutamine synthetase type III, GlnN (EC 6.3.1.2)	4.08E+02	0.00E+00	1.18E+03	0	0	0	0
Glutamine synthetase, clostridia type (EC 6.3.1.2)	0.00E+00	3.60E+04	0	0	0	0	0
Nitrogen regulation protein NR(I)	3.19E+03	4.35E+03	4.66E+03	0	0	0	0
Nitrogen regulatory protein P-II	6.03E+02	7.92E+02	2.32E+03	4.89E+02	1.34E+03	3.42E+02	0.00E+00
Nitrogen regulatory protein P-II, glnK	0.00E+00	1.40E+03	0	0	0	0	0
[Protein-PII] uridylyltransferase (EC 2.7.7.59)	4.45E+03	0.00E+00	8.86E+03	0	0	0	0

Subsystem: Cyanate hydrolysis	Number of aligned bases (b)

	DAL	GAL	BLA	LADP5	LADP6	LA	MUP
Carbonic anhydrase (EC 4.2.1.1)	1.59E+03	2.37E+03	6.03E+03	0.00E+00	4.68E+02	0	0
Cyanate ABC transporter, ATP-binding protein	0	0	5.01E+02	0	0	0	0
Cyanate ABC transporter, permease protein	0	0	3.21E+02	0	0	0	0
Cyanate ABC transporter, substrate binding protein	4.44E+02	0.00E+00	1.49E+03	0	0	0	0
Cyanate hydratase (EC 4.2.1.104)	3.90E+02	0	0	0	0	0	0
Cyanate transport protein CynX	0	0	2.97E+02	0	0	0	0
Cyn operon transcriptional activator	0	0	5.43E+02	0	0	0	0
Subsystem: Denitrification	Number of aligned bases (b)						
	DAL	GAL	BLA	LADP5	LADP6	LA	MUP
Copper-containing nitrite reductase (EC 1.7.2.1)	1.77E+03	0.00E+00	1.28E+03	0.00E+00	6.09E+02	2.13E+02	0.00E+00
Nitric oxide -responding transcriptional regulator NnrR (Crp/Fnr family)	2.19E+02	0.00E+00	0.00E+00	2.52E+02	0.00E+00	0.00E+00	0.00E+00
Nitric oxide reductase activation protein NorE	0.00E+00	2.90E+03	0.00E+00	0.00E+00	0.00E+00	0.00E+00	2.22E+02
Nitric oxide reductase activation protein NorQ	0.00E+00	6.51E+03	2.85E+02	0.00E+00	5.70E+02	2.28E+02	0.00E+00
Nitric-oxide reductase (EC 1.7.99.7), quinol-dependent	4.52E+03	0.00E+00	3.73E+03	8.25E+02	5.07E+02	0.00E+00	0.00E+00
Nitric-oxide reductase subunit B (EC 1.7.99.7)	0.00E+00	8.76E+02	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00
Nitrite reductase accessory protein NirV	0.00E+00	4.80E+02	0.00E+00	0.00E+00	0.00E+00	0.00E+00	4.98E+02
Nitrous oxide reductase maturation periplasmic protein NosX	0.00E+00	4.50E+03	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00
Nitrous oxide reductase maturation protein NosD	3.27E+02	2.28E+02	2.19E+02	0.00E+00	0.00E+00	2.31E+02	0.00E+00
Nitrous oxide reductase maturation protein NosF (ATPase)	2.04E+02	0.00E+00	3.99E+02	6.19E+02	6.88E+02	7.14E+02	0.00E+00
Nitrous oxide reductase maturation protein NosR	0.00E+00	0.00E+00	6.12E+02	0.00E+00	0.00E+00	0.00E+00	0.00E+00
Nitrous oxide reductase maturation protein, outer-membrane lipoprotein NosL	0.00E+00	2.67E+02	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00
Nitrous-oxide reductase (EC 1.7.99.6)	0.00E+00	0.00E+00	5.67E+02	1.68E+02	3.75E+02	2.04E+02	0.00E+00
NnrS protein involved in response to NO	2.22E+02	0.00E+00	1.42E+03	0.00E+00	0.00E+00	0.00E+00	0.00E+00
NnrU family protein, required for expression of nitric oxide and nitrite reductases (Nir and Nor)	0.00E+00	4.82E+03	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00
Subsystem: Dissimilatory nitrite reductase	Number of aligned bases (b)						
	DAL	GAL	BLA	LADP5	LADP6	LA	MUP
Assimilatory nitrate reductase large subunit (EC	2.61E+03	0.00E+00	3.30E+03	0	0	0	0
Cytochrome c-type heme lyase subunit nrfE, nitrite reductase complex assembly	0	0	6.93E+02	0	0	0	0
Cytochrome c-type heme lyase subunit nrfG, nitrite reductase complex assembly	0	0	4.38E+02	0	0	0	0
Cytochrome c-type protein NapC	9.12E+02	0.00E+00	2.85E+02	0	0	0	0

Cytochrome c552 precursor (EC 1.7.2.2)	4.74E+02	0.00E+00	8.07E+02	0	0	0	0
Ferredoxin--nitrite reductase (EC 1.7.7.1)	3.54E+02	0.00E+00	1.89E+03	2.94E+02	1.18E+03	1.76E+03	8.01E+02
Ferredoxin-type protein NapF (periplasmic nitrate reductase)	0	0	2.22E+02	0	0	0	0
Ferredoxin-type protein NapG (periplasmic nitrate reductase)	2.40E+02	0	0	0	0	0	0
Nitrate ABC transporter, ATP-binding protein	0	0	2.05E+03	0	0	0	0
Nitrate ABC transporter, nitrate-binding protein	1.13E+03	0.00E+00	2.73E+03	0	0	0	0
Nitrate ABC transporter, permease protein	5.10E+02	0.00E+00	7.02E+02	0	0	0	0
Nitrate reductase cytochrome c550-type subunit	0	0	1.77E+02	0	0	0	0
Nitrate/nitrite response regulator protein	2.34E+02	0.00E+00	3.96E+02	0	0	0	0
Nitrate/nitrite sensor protein (EC 2.7.3.-)	0	0	2.50E+03	0	0	0	0
Nitrate/nitrite transporter	8.77E+03	0.00E+00	1.87E+04	2.07E+02	0	0	0
Nitrite reductase probable [NAD(P)H] subunit (EC 1.7.1.4)	0	0	2.61E+02	2.04E+02	5.23E+02	2.01E+03	3.63E+02
Nitrite reductase [NAD(P)H] large subunit (EC 1.7.1.4)	2.81E+03	0.00E+00	8.56E+03	0	0	0	0
Nitrite reductase [NAD(P)H] small subunit (EC 1.7.1.4)	2.19E+02	0.00E+00	9.81E+02	0	0	0	0
Nitrite transporter from formate/nitrite family	2.16E+02	0.00E+00	6.48E+02	0	0	0	0
NrfD protein	2.22E+02	0.00E+00	3.24E+02	0	0	0	0
Periplasmic nitrate reductase precursor (EC 1.7.99.4)	1.47E+03	0.00E+00	4.42E+03	0	0	0	0
Polyferredoxin NapH (periplasmic nitrate reductase)	0	0	2.34E+02	0	0	0	0
Respiratory nitrate reductase alpha chain (EC 1.7.99.4)	3.79E+03	0.00E+00	1.06E+04	1.20E+03	2.43E+03	2.67E+03	4.56E+02
Respiratory nitrate reductase beta chain (EC 1.7.99.4)	4.25E+03	0.00E+00	2.99E+03	0.00E+00	1.26E+03	5.88E+02	8.82E+02
Respiratory nitrate reductase delta chain (EC 1.7.99.4)	3.99E+02	0	0	0	0	0	0
Respiratory nitrate reductase gamma chain (EC 1.7.99.4)	5.25E+02	0.00E+00	7.11E+02	0	0	0	0
Response regulator NasT	5.79E+02	0	0	0	0	0	0
Subsystem: Nitric oxide synthase	Number of aligned bases (b)						
	DAL	GAL	BLA	LADP5	LADP6	LA	MUP
4Fe-4S ferredoxin, nitrogenase-associated	2.22E+02	0.00E+00	2.22E+02	0	0	0	0
Cysteine desulfurase (EC 2.8.1.7), NifS subfamily	2.25E+02	0.00E+00	1.36E+03	0	0	0	0
Homocitrate synthase (EC 2.3.3.14)	2.22E+02	0.00E+00	7.68E+02	2.22E+02	0	0	0

NifT protein	0	0	2.19E+02	0	0	0	0
NifX-associated protein	3.69E+02	0	0	0	0	0	0
NifZ protein	2.40E+02	0	0	0	0	0	0
Nitrogenase (molybdenum-iron) alpha chain (EC 1.18.6.1)	0	0	7.44E+02	0	0	0	0
Nitrogenase (molybdenum-iron) beta chain (EC 1.18.6.1)	9.96E+02	0.00E+00	1.65E+03	0	0	2.55E+02	0.00E+00
Nitrogenase (molybdenum-iron) reductase and maturation protein NifH	2.70E+02	0.00E+00	8.38E+02	0	0	0	0
Nitrogenase (molybdenum-iron)-specific transcriptional regulator NifA	0	0	2.18E+03	0	0	0	0
Nitrogenase FeMo-cofactor carrier protein NifX	0	0	1.80E+02	0	0	0	0
Nitrogenase FeMo-cofactor scaffold and assembly protein NifE	4.83E+02	0.00E+00	1.45E+03	0	0	0	0
Nitrogenase FeMo-cofactor scaffold and assembly protein NifN	0	0	2.13E+02	0.00E+00	5.49E+02	0	0
Nitrogenase FeMo-cofactor synthesis FeS core scaffold and assembly protein NifB	7.44E+02	0.00E+00	2.29E+03	0	0	2.07E+02	0.00E+00
Nitrogenase FeMo-cofactor synthesis molybdenum delivery protein NifQ	2.58E+02	0	0	0	0	0	0
Nitrogenase stabilizing/protective protein NifW	0	0	7.23E+02	0	0	0	0

Subsystem:Nitrogen Fixation with NifL	Number of aligned bases (b)						
	DAL	GAL	BLA	LADP5	LADP6	LA	MUP
4Fe-4S ferredoxin, nitrogenase-associated	2.22E+02	0.00E+00	2.22E+02	0	0	0	0
Cysteine desulfurase (EC 2.8.1.7)	5.88E+03	0.00E+00	5.51E+03	1.15E+03	2.81E+03	1.60E+03	2.73E+02
Cysteine desulfurase (EC 2.8.1.7), NifS subfamily	2.25E+02	0.00E+00	1.36E+03	0	0	0	0
Homocitrate synthase (EC 2.3.3.14)	2.22E+02	0.00E+00	7.68E+02	2.22E+02	0	0	0
NifT protein	0	0	2.19E+02	0	0	0	0
NifX-associated protein	3.69E+02	0	0	0	0	0	0
NifZ protein	2.40E+02	0	0	0	0	0	0
Nitrogenase (molybdenum-iron) alpha chain (EC 1.18.6.1)	0	0	7.44E+02	0	0	0	0
Nitrogenase (molybdenum-iron) beta chain (EC 1.18.6.1)	9.96E+02	0.00E+00	1.65E+03	0	0	2.55E+02	0.00E+00
Nitrogenase (molybdenum-iron) reductase and maturation protein NifH	2.70E+02	0.00E+00	8.38E+02	0	0	0	0
Nitrogenase (molybdenum-iron)-specific transcriptional regulator NifA	0	0	2.18E+03	0	0	0	0
Nitrogenase FeMo-cofactor carrier protein NifX	0	0	1.80E+02	0	0	0	0

Nitrogenase FeMo-cofactor scaffold and assembly protein NifE	4.83E+02	0.00E+00	1.45E+03	0	0	0	0
Nitrogenase FeMo-cofactor scaffold and assembly protein NifN	0	0	2.13E+02	0.00E+00	5.49E+02	0	0
Nitrogenase FeMo-cofactor synthesis FeS core scaffold and assembly protein NifB	7.44E+02	0.00E+00	2.29E+03	0	0	2.07E+02	0.00E+00
Nitrogenase FeMo-cofactor synthesis molybdenum delivery protein NifQ	2.58E+02	0	0	0	0	0	0
Nitrogenase stabilizing/protective protein NifW	0	0	7.23E+02	0	0	0	0

Subsystem:Nitrosative stress	Number of aligned bases (b)						
	DAL	GAL	BLA	LADP5	LADP6	LA	MUP
Anaerobic nitric oxide reductase flavorubredoxin	2.94E+02	2.67E+03	5.58E+02	0	0	0	0
Ferredoxin 3 fused to uncharacterized domain	0.00E+00	2.75E+03	0	0	0	0	0
Functional role page for Anaerobic nitric oxide reductase transcription regulator NorR	0	0	2.08E+02	0	0	0	0
Hcp transcriptional regulator HcpR (Crp/Fnr family)	0.00E+00	1.20E+04	0	0	0	0	0
Hydroxylamine reductase (EC 1.7.-.-)	0	0	7.89E+02	0	0	0	0
NADH oxidoreductase hcr (EC 1.-.-.-)	2.10E+02	2.85E+02	0	0	0	0	0
Nitric-oxide reductase (EC 1.7.99.7), quinol-dependent	4.52E+03	0.00E+00	3.73E+03	8.25E+02	5.07E+02	4.77E+02	4.98E+02
Nitrite-sensitive transcriptional repressor NsrR	0.00E+00	1.11E+04	0	0	0	0	0
NnrS protein involved in response to NO	2.22E+02	0.00E+00	1.42E+03	0	0	0	0

Appendix 7: Protein families predicted in subsystems of Sulfur metabolism

Subsystem:Alkanesulfonate assimilation	**Number of aligned bases (b)**						
	DAL	**GAL**	**BLA**	**LADP5**	**LADP6**	**LA**	**MUP**
ABC-type nitrate/sulfonate/bicarbonate transport system, ATPase component	2646	9636	4041	0	0	0	0
ABC-type nitrate/sulfonate/bicarbonate transport system, permease component	813	0	924	306	630	267	303
ABC-type nitrate/sulfonate/bicarbonate transport systems, periplasmic components	3588	0	6734	0	0	0	0
Alkanesulfonate monooxygenase (EC 1.14.14.5)	1661	570	3858	0	0	0	0
Alkanesulfonates ABC transporter ATP-binding protein	936	5971	2745	0	0	0	0
Alkanesulfonates transport system permease protein	1665	0	6255	0	0	0	0
Alkanesulfonates-binding protein	3146	0	6431	0	0	0	0
Alpha-ketoglutarate-dependent taurine dioxygenase (EC 1.14.11.17)	2316	0	4104	0	0	0	0
Arylsulfatase (EC 3.1.6.1)	3456	0	7262	222	252	234	0
FMN reductase (EC 1.5.1.29)	0	0	432	0	0	0	0
Organosulfonate utilization protein SsuF	156	0	0	0	0	0	0
probable dibenzothiophene desulfurization enzyme	0	0	2650	0	0	0	0
Putative arylsulfatase regulatory protein	0	1905	0	0	0	0	0
sulfonate monooxygenase	0	8101	276	0	0	0	0

Subsystem:Alkanesulfonates Utilization	**Number of aligned bases (b)**						
	DAL	**GAL**	**BLA**	**LADP5**	**LADP6**	**LA**	**MUP**
Alkanesulfonate monooxygenase (EC 1.14.14.5)	1661	570	3858	0	0	0	0
Alkanesulfonates ABC transporter ATP-binding protein	936	5971	2745	0	0	0	0
Alkanesulfonates transport system permease protein	1665	0	6255	0	0	0	0
Alkanesulfonates-binding protein	3146	0	6431	0	0	0	0
FMN reductase (EC 1.5.1.29)	0	0	432	0	0	0	0

Subsystem:At5g37530 (CsdL protein family)	**Number of aligned bases (b)**						
	DAL	**GAL**	**BLA**	**LADP5**	**LADP6**	**LA**	**MUP**
Aspartyl-tRNA synthetase (EC 6.1.1.12)	5326	0	8052	7235	7350	6018	1098
CsdL (EC-YgdL) protein of the HesA/MoeB/ThiF family, part of the CsdA-E-L sulfur transfer pathway	276	0	0	0	681	0	0
Cysteine desulfurase (EC 2.8.1.7)	5882	0	5508	1152	2811	1598	273
Cysteine desulfurase (EC 2.8.1.7), IscS subfamily	1365	0	360	0	0	285	0
Cysteine desulfurase (EC 2.8.1.7), NifS subfamily	225	0	1362	0	0	0	0
Cysteine desulfurase (EC 2.8.1.7), SufS subfamily	2520	0	3950	207	2358	1365	2457
Cysteine desulfurase CsdA-CsdE (EC 2.8.1.7), main protein CsdA	702	0	573	0	0	0	0
Cysteine desulfurase CsdA-CsdE, sulfur acceptor protein CsdE	0	0	267	0	0	0	0
Iron-sulfur cluster assembly protein SufB	2859	0	4983	2100	7451	3786	3735
Iron-sulfur cluster regulator IscR	861	219	991	0	0	0	0
tRNA (5-methylaminomethyl-2-thiouridylate)-methyltransferase (EC 2.1.1.61)	3949	0	5254	2175	486	1224	159
tRNA-specific 2-thiouridylase MnmA	0	0	561	0	0	0	0

Subsystem:Bacterial pathways for dimethylsulfoniopropionate and acrylate catabolism	**Number of aligned bases (b)**						
	DAL	**GAL**	**BLA**	**LADP5**	**LADP6**	**LA**	**MUP**
3-methylmercaptopropionyl-CoA dehydrogenase (DmdC)	3135	0	7761	0	0	0	0
3-methylmercaptopropionyl-CoA ligase (DmdB)	1059	0	3804	0	0	0	0
Acryloyl-CoA reductase AcuI/YhdH (EC 1.3.1.84)	321	0	1887	0	0	0	0

Subsystem:CFE Sulfur Oxidation	**Number of aligned bases (b)**						
	DAL	**GAL**	**BLA**	**LADP5**	**LADP6**	**LA**	**MUP**
Adenylylsulfate reductase alpha-subunit (EC 1.8.99.2)	0	0	1004	0	0	0	0
Anaerobic dimethyl sulfoxide reductase chain B (EC 1.8.5.3)	783	0	345	0	0	0	0
Blr3520 protein homolog, hypothetical protein	0	0	942	0	0	0	0
Cytochrome c-type biogenesis protein CcdA (DsbD analog)	1119	0	3864	0	411	0	0
heterodisulfide reductase, subunit A/methylviologen reducing hydrogenase, subunit delta	0	23905	0	0	0	0	0
Lipocalin-related protein and Bos/Can/Equ allergen	270	0	210	0	0	0	0
Sulfide dehydrogenase [flavocytochrome C] flavoprotein chain precursor (EC 1.8.2.-)	489	0	1242	0	0	0	0
Sulfur oxidation molybdopterin C protein	702	0	552	0	0	0	0
sulfur oxidation protein SoxA	0	0	1458	0	0	0	0
Sulfur oxidation protein SoxB	1140	0	2478	0	0	0	0
Sulfur oxidation protein SoxX	237	0	0	0	0	0	0
Sulfur oxidation protein SoxY	0	0	519	0	0	0	0
Sulfur oxidation protein SoxZ	240	0	783	0	0	0	0

Subsystem:Dimethylsulfoniopropionate (DMSP) mineralization	**Number of aligned bases (b)**						
	DAL	**GAL**	**BLA**	**LADP5**	**LADP6**	**LA**	**MUP**
3-methylmercaptopropionyl-CoA dehydrogenase (DmdC)	3135	0	7761	0	0	0	0

3-methylmercaptopropionyl-CoA ligase (DmdB)	1059	0	3804	0	0	0	0

Subsystem:Dimethylsulfoniopropionate (DMSP) mineralization, gjo	**Number of aligned bases (b)**						
	DAL	**GAL**	**BLA**	**LADP5**	**LADP6**	**LA**	**MUP**
3-methylmercaptopropionyl-CoA dehydrogenase (DmdC)	3135	0	7761	0	0	0	0
3-methylmercaptopropionyl-CoA ligase (DmdB)	1059	0	3804	0	0	0	0
Acryloyl-CoA reductase AcuI/YhdH (EC 1.3.1.84)	321	0	1887	0	0	0	0
Cysteine desulfurase (EC 2.8.1.7)	5882	0	5508	1152	2811	1598	273

Subsystem:Dimethylsulfoniopropionate (DMSP) mineralization, WBW	**Number of aligned bases (b)**						
	DAL	**GAL**	**BLA**	**LADP5**	**LADP6**	**LA**	**MUP**
3-methylmercaptopropionyl-CoA dehydrogenase (DmdC)	3135	0	7761	0	0	0	0
3-methylmercaptopropionyl-CoA ligase (DmdB)	1059	0	3804	0	0	0	0
Acryloyl-CoA reductase AcuI/YhdH (EC 1.3.1.84)	321	0	1887	0	0	0	0
Cysteine desulfurase (EC 2.8.1.7)	5882	0	5508	1152	2811	1598	273

Subsystem:Galactosylceramide and Sulfatide metabolism	**Number of aligned bases (b)**						
	DAL	**GAL**	**BLA**	**LADP5**	**LADP6**	**LA**	**MUP**
Arylsulfatase (EC 3.1.6.1)	3456	0	7262	222	252	234	0
Beta-galactosidase (EC 3.2.1.23)	6714	32684	6045	0	0	0	0
Neuraminidase NanP	930	0	264	0	0	0	0
Sialidase (EC 3.2.1.18)	7752	0	4509	0	0	0	0

Subsystem:Inorganic Sulfur Assimilation	**Number of aligned bases (b)**						
	DAL	**GAL**	**BLA**	**LADP5**	**LADP6**	**LA**	**MUP**
3'(2'),5'-bisphosphate nucleotidase (EC 3.1.3.7)	0	0	1065	0	0	0	0
Adenylylsulfate kinase (EC 2.7.1.25)	951	0	2732	465	3684	1524	1356
Adenylylsulfate reductase alpha-subunit (EC 1.8.99.2)	0	0	1004	0	0	0	0
Ferredoxin	597	0	1065	1337	981	2831	954
Ferredoxin--NADP(+) reductase (EC 1.18.1.2)	2415	0	5032	471	387	231	0
Phosphoadenylyl-sulfate reductase [thioredoxin] (EC 1.8.4.8)	861	0	2091	417	648	276	261
Sulfate adenylyltransferase subunit 1 (EC 2.7.7.4)	891	0	1068	0	0	0	0
Sulfate adenylyltransferase subunit 2 (EC 2.7.7.4)	1326	0	1598	1464	450	213	1716
Sulfate and thiosulfate binding protein CysP	1698	0	3866	0	0	0	0
Sulfate and thiosulfate import ATP-binding protein CysA (EC 3.6.3.25)	837	0	1605	1443	1512	1310	0
Sulfate permease, Trk-type	0	0	0	0	327	0	0
Sulfate transport system permease protein CysT	2355	0	1657	0	0	0	0
Sulfate transport system permease protein CysW	501	0	1354	0	0	0	0
Sulfate-binding protein Sbp	942	0	432	0	0	0	0
Sulfite reductase [NADPH] flavoprotein alpha-component (EC 1.8.1.2)	477	0	924	0	2772	837	0
Sulfite reductase [NADPH] hemoprotein beta-component (EC 1.8.1.2)	3890	0	8547	219	2943	693	0

Subsystem:L-Cystine Uptake and Metabolism	**Number of aligned bases (b)**						
	DAL	**GAL**	**BLA**	**LADP5**	**LADP6**	**LA**	**MUP**
Protein similar to glutamate synthase [NADPH] small chain, clustered with sulfite reductase	0	7159	0	0	0	0	0
Sulfite reductase, dissimilatory-type gamma subunit (EC 1.8.99.3)	0	1596	0	0	0	0	0
Sulfite reduction-associated complex DsrMKJOP iron-sulfur protein DsrO (=HmeA)	0	7749	0	0	0	0	0
Sulfite reduction-associated complex DsrMKJOP multiheme protein DsrJ (=HmeF)	0	12051	0	0	0	0	0
Sulfite reduction-associated complex DsrMKJOP protein DsrM (= HmeC)	0	5283	0	0	0	0	0
Sulfite reduction-associated complex DsrMKJOP protein DsrP (= HmeB)	0	12184	0	0	0	0	0
tRNA 2-thiouridine synthesizing protein E (EC 2.8.1.-)	0	516	0	0	0	0	0

Subsystem:Sulfate reduction-associated complexes	**Number of aligned bases (b)**						
	DAL	**GAL**	**BLA**	**LADP5**	**LADP6**	**LA**	**MUP**
Protein similar to glutamate synthase [NADPH] small chain, clustered with sulfite reductase	0	7159	0	0	0	0	0
Sulfite reductase, dissimilatory-type gamma subunit (EC 1.8.99.3)	0	1596	0	0	0	0	0
Sulfite reduction-associated complex DsrMKJOP iron-sulfur protein DsrO (=HmeA)	0	7749	0	0	0	0	0
Sulfite reduction-associated complex DsrMKJOP multiheme protein DsrJ (=HmeF)	0	12051	0	0	0	0	0
Sulfite reduction-associated complex DsrMKJOP protein DsrM (= HmeC)	0	5283	0	0	0	0	0
Sulfite reduction-associated complex DsrMKJOP protein DsrP (= HmeB)	0	12184	0	0	0	0	0
tRNA 2-thiouridine synthesizing protein E (EC 2.8.1.-)	0	516	0	0	0	0	0

Subsystem:Sulfur oxidation	**Number of aligned bases (b)**						
	DAL	**GAL**	**BLA**	**LADP5**	**LADP6**	**LA**	**MUP**
Blr3520 protein homolog, hypothetical protein	0	0	942	0	0	0	0
Cytochrome c-type biogenesis protein CcdA (DsbD analog)	1119	0	3864	0	411	0	0
Lipocalin-related protein and Bos/Can/Equ allergen	270	0	210	0	0	0	0
Sulfide dehydrogenase [flavocytochrome C] flavoprotein chain precursor (EC 1.8.2.-)	489	0	1242	0	0	0	0
Sulfur oxidation molybdopterin C protein	702	0	552	0	0	0	0
sulfur oxidation protein SoxA	0	0	1458	0	0	0	0
Sulfur oxidation protein SoxB	1140	0	2478	0	0	0	0
Sulfur oxidation protein SoxX	237	0	0	0	0	0	0
Sulfur oxidation protein SoxY	0	0	519	0	0	0	0
Sulfur oxidation protein SoxZ	240	0	783	0	0	0	0

Subsystem:Taurine Utilization	**Number of aligned bases (b)**						
	DAL	**GAL**	**BLA**	**LADP5**	**LADP6**	**LA**	**MUP**
Alpha-ketoglutarate-dependent taurine dioxygenase (EC 1.14.11.17)	2316	0	4104	0	0	0	0
Taurine transport ATP-binding protein TauB	0	0	1314	0	0	0	0
Taurine transport system permease protein TauC	489	0	369	0	0	0	0
Taurine-binding periplasmic protein TauA	0	0	525	0	0	0	0

Subsystem:Thioredoxin-disulfide reductase	**Number of aligned bases (b)**						
	DAL	**GAL**	**BLA**	**LADP5**	**LADP6**	**LA**	**MUP**
Alkyl hydroperoxide reductase protein C (EC 1.6.4.-)	969	0	180	0	0	0	0
Alkyl hydroperoxide reductase protein F (EC 1.6.4.-)	2049	0	2439	0	2355	0	0
Alkyl hydroperoxide reductase subunit C-like protein	1290	0	2679	816	669	435	0
Hydrogen peroxide-inducible genes activator	2001	0	3783	453	312	0	507
Thiol peroxidase, Bcp-type (EC 1.11.1.15)	279	0	0	0	0	0	0
Thiol peroxidase, Tpx-type (EC 1.11.1.15)	0	0	927	0	312	0	0
Thioredoxin reductase (EC 1.8.1.9)	6336	5406	16482	4077	12068	5820	1032

Subsystem:Utilization of glutathione as a sulphur source	**Number of aligned bases (b)**						
	DAL	**GAL**	**BLA**	**LADP5**	**LADP6**	**LA**	**MUP**
Gamma-glutamyltranspeptidase (EC 2.3.2.2)	4646	0	11822	0	0	0	537
Putative glutathione transporter, ATP-binding component	450	0	1689	0	0	0	0
Putative glutathione transporter, permease component	1221	0	1050	0	0	0	0
Putative glutathione transporter,solute-binding component	0	0	618	0	0	0	0
Putative gluthatione transporter,solute-binding component	0	0	465	0	0	0	0

Appendix 8: Protein families predicted in subsystems of Phosphorous metabolism

Subsystem: Alkylphosphonate utilization	**Number of aligned bases (b)**						
	DAL	**GAL**	**BLA**	**LADP5**	**LADP6**	**LA**	**MUP**
Alkylphosphonate utilization operon protein PhnA	0	0	282	0	0	0	0
Metal-dependent hydrolase involved in phosphonate metabolism	453	1881	3477	0	0	0	0
PhnB protein	354	0	270	0	0	0	0
PhnG protein	210	0	0	0	0	0	0
PhnH protein	0	594	658	0	0	0	0
PhnI protein	1074	0	2940	0	0	0	0
PhnJ protein	465	0	1350	219	0	0	0
Phosphonates transport ATP-binding protein PhnK	0	0	1170	0	0	0	0
Phosphonates transport ATP-binding protein PhnL	0	0	861	0	0	0	0
Protein RcsF	435	0	324	0	0	0	0
Transcriptional regulator PhnF	756	0	1155	0	0	0	0

Subsystem: High affinity phosphate transporter and control of PHO regulon	**Number of aligned bases (b)**						
	DAL	**GAL**	**BLA**	**LADP5**	**LADP6**	**LA**	**MUP**
Alkaline phosphatase synthesis transcriptional regulatory protein PhoP	0	0	0	0	207	0	0
Phosphate ABC transporter, periplasmic phosphate-binding protein PstS (TC 3.A.1.7.1)	2727	0	5466	243	834	267	0
Phosphate regulon sensor protein PhoR (SphS) (EC 2.7.13.3)	4257	0	7472	0	1125	0	0
Phosphate regulon transcriptional regulatory protein PhoB (SphR)	1959	0	3033	1392	831	204	942
Phosphate transport ATP-binding protein PstB (TC 3.A.1.7.1)	243	0	2409	519	1878	2430	2331
Phosphate transport system permease protein PstA (TC 3.A.1.7.1)	3856	0	912	1050	1464	4593	2379
Phosphate transport system permease protein PstC (TC 3.A.1.7.1)	2037	0	2379	1656	3870	2505	648
Phosphate transport system regulatory protein PhoU	684	0	2955	2964	2878	4221	585
Polyphosphate kinase (EC 2.7.4.1)	9450	0	11666	2001	4359	4760	774
Putative periplasmic phosphate-binding protein PstS (Halobacteriales type)	0	0	0	207	210	489	1965

Subsystem: P uptake (cyanobacteria)	**Number of aligned bases (b)**						
	DAL	**GAL**	**BLA**	**LADP5**	**LADP6**	**LA**	**MUP**
Alkaline phosphatase (EC 3.1.3.1)	1878	0	5259	1077	2229	1947	366
Phosphate ABC transporter, periplasmic phosphate-binding protein PstS (TC 3.A.1.7.1)	2727	0	5466	243	834	267	0
Phosphate transport ATP-binding protein PstB (TC 3.A.1.7.1)	243	0	2409	519	1878	2430	2331
Phosphate transport system permease protein PstA (TC 3.A.1.7.1)	3856	0	912	1050	1464	4593	2379
Phosphate transport system permease protein PstC (TC 3.A.1.7.1)	2037	0	2379	1656	3870	2505	648

Subsystem: Phosphate metabolism	**Number of aligned bases (b)**						
	DAL	**GAL**	**BLA**	**LADP5**	**LADP6**	**LA**	**MUP**
1-acyl-sn-glycerol-3-phosphate acyltransferase (EC 2.3.1.51)	4374	0	6169	0	0	0	0
Alkaline phosphatase (EC 3.1.3.1)	1878	0	5259	1077	2229	1947	366
Alkaline phosphatase synthesis transcriptional regulatory protein PhoP	0	0	0	0	207	0	0
Apolipoprotein N-acyltransferase (EC 2.3.1.-)	4275	0	5604	0	0	0	0
Exopolyphosphatase (EC 3.6.1.11)	1641	0	3732	0	0	0	0
Inorganic pyrophospatase PpaX (EC 3.1.3.18)	0	0	0	0	1415	0	0
Inorganic pyrophosphatase (EC 3.6.1.1)	1599	0	1578	1908	3792	1032	348
Integral membrane protein YggT, involved in response to extracytoplasmic stress (osmotic shock)	1548	0	4042	0	0	0	0
Low-affinity inorganic phosphate transporter	792	0	3668	0	0	0	0
Magnesium and cobalt efflux protein CorC	5390	0	4033	870	1095	1065	1302
Manganese-dependent inorganic pyrophosphatase (EC 3.6.1.1)	225	0	225	0	5559	0	0

Metal-dependent hydrolase YbeY, involved in rRNA and/or ribosome maturation and assembly	915	0	1332	0	0	0	0
NAD(P) transhydrogenase alpha subunit (EC 1.6.1.2)	2688	0	7262	1377	1215	273	0
NAD(P) transhydrogenase subunit beta (EC 1.6.1.2)	4560	0	6363	4595	3882	4847	771
Phosphatase, Ppx/GppA family	381	0	411	0	0	0	0
Phosphate ABC transporter, periplasmic phosphate-binding protein PstS (TC 3.A.1.7.1)	2727	0	5466	243	834	267	0
Phosphate regulon sensor protein PhoR (SphS) (EC 2.7.13.3)	4257	0	7472	0	1125	0	0
Phosphate regulon transcriptional regulatory protein PhoB (SphR)	1959	0	3033	1392	831	204	942
Phosphate starvation-inducible protein PhoH, predicted ATPase	1011	0	2443	0	1218	747	252
Phosphate transport ATP-binding protein PstB (TC 3.A.1.7.1)	243	0	2409	519	1878	2430	2331
Phosphate transport regulator (distant homolog of PhoU)	927	0	915	0	0	0	0
Phosphate transport system permease protein PstA (TC 3.A.1.7.1)	3856	0	912	1050	1464	4593	2379
Phosphate transport system permease protein PstC (TC 3.A.1.7.1)	2037	0	2379	1656	3870	2505	648
Phosphate transport system regulatory protein PhoU	684	0	2955	2964	2878	4221	585
Polyphosphate kinase (EC 2.7.4.1)	9450	0	11666	2001	4359	4760	774
Predicted ATPase related to phosphate starvation-inducible protein PhoH	1821	0	855	0	1326	0	0
Probable low-affinity inorganic phosphate transporter	2471	0	6066	477	2108	3336	3006
Pyrophosphate-energized proton pump (EC 3.6.1.1)	3864	0	7664	249	2076	355	0
Pyrophosphate-specific outer membrane porin OprO	480	0	0	0	0	0	0
response regulator in two-component regulatory system with PhoQ	231	0	2160	0	0	0	0
secreted alkaline phosphatase	1560	0	7967	0	0	0	0
Sodium-dependent phosphate transporter	2040	0	1347	0	633	0	0
Soluble pyridine nucleotide transhydrogenase (EC 1.6.1.1)	1125	0	486	0	0	0	0

Subsystem: Phosphoenolpyruvate phosphomutase	**Number of aligned bases (b)**						
	DAL	**GAL**	**BLA**	**LADP5**	**LADP6**	**LA**	**MUP**
2-aminoethylphosphonate	552	0	3314	0	0	0	0
Phosphonopyruvate decarboxylase (EC 4.1.1.82)	558	0	0	0	0	0	0

Subsystem: Phosphonate metabolism	**Number of aligned bases (b)**						
	DAL	**GAL**	**BLA**	**LADP5**	**LADP6**	**LA**	**MUP**
2-aminoethylphosphonate ABC transporter ATP-binding protein (TC 3.A.1.9.1)	0	453	453	0	0	0	0
2-aminoethylphosphonate ABC transporter permease protein I (TC 3.A.1.9.1)	0	210	0	0	0	0	0
2-aminoethylphosphonate	5520	0	3314	0	0	0	0
Phosphonoacetate hydrolase (EC 3.11.1.2)	0	174	0	0	0	0	0
Phosphonopyruvate decarboxylase (EC 4.1.1.82)	558	0	0	0	0	0	0

Appendix 9: Number of aligned bases assigned in kilo base (kb) to the five nutrient metabolism subsystems of the most abundant taxonomic Order

		Nitrogen	Sulfur	Phosphorus	Carbohydrate	Potassium
DAL						
	Rhizobiales	17.8 kb	31.6 kb	12.5 kb	143.7 kb	15.6 kb
	Burholderiales	25.7 kb	26.5 kb	17.9 kb	192.8 kb	13.8 kb
	Pseudomonadales	3.8 kb	6.5 kb	13.4 kb	41.8 kb	3.9 kb
	Entrobacteriales	7.5 kb	6.4 kb	2.90 kb	55.0 kb	4.1 kb
	Propionibacteriales	24.2kb	13.9 kb	11.6 kb	163.9 kb	10.3 kb
	Others	26.0 kb	28.1 kb	18.6 kb	239.8 kb	
GAL						
	Burholderiales	66.8 kb	51.4 kb	0.66 kb	527.3 kb	4.5 kb
	Rhizobiales	46.8 kb	51.2 kb	0 kb	293.5 kb	0 kb
	Propionibacteriales	4.2 kb	6.6 kb	0.2 kb	111.9 kb	0 kb
	Entrobacteriales	1.9 kb	1.7 kb	0.2 kb	17.7 kb	0 kb
	Corynebacteriales	0.5 kb	0.25 kb	0 kb	7.9 kb	0 kb
	Others	44.8 kb	26.85 kb	1.8 kb	391.7 kb	
BLA						
	Rhizobiales	69.8 kb	78.9 kb	50.1 kb	458.8 kb	59.0 kb
	Burholderiales	28.7 kb	31.5 kb	18.0 kb	198.1 kb	15.9 kb
	Pseudomonadales	0.2 kb	0.2 kb	0.2 kb	1.0 kb	0 kb
	Entrobacteriales	25.1 kb	12.0 kb	7.0 kb	156.0 kb	14.1 kb
	Propionibacteriales	20.6 kb	16.2 kb	14.8 kb	180.8 kb	11.7 kb
	Others	70.6 kb	59.2 kb	41.9 kb	445.3 kb	
LADP5						
	Haloferacales	0.2 kb	1.4 kb	1.6 kb	14.9 kb	1.6 kb
	Halobacteriales	0.9 kb	1.2 kb	1.2 kb	7.5 kb	0.2 kb
	Bacteroidetes+	4.4 kb	13.1 kb	8.8 kb	95.8 kb	5.1 kb
	Others	7.5 kb	9.3 kb	23.4 kb	111.8 kb	
LADP6						
	Haloferacales	1.1 kb	6.0 kb	3.1 kb	17.5 kb	2.2 kb
	Halobacteriales	3.0 kb	1.6 kb	1.8 kb	21.6 kb	4.0 kb
	Bacteroidetes+	6.7 kb	11.3 kb	7.5 kb	128.3 kb	10.0 kb
	Others	22.2 kb	33.1 kb	59.6 kb	262.6 kb	
LA						
	Haloferacales	1.3 kb	1.0 kb	1.8 kb	12.4 kb	5.6 kb
	Halobacteriales	5.4 kb	2.4 kb	5.6 kb	22.4 kb	2.7 kb
	Bacteroidetes+	7.4 kb	14.4 kb	8.3 kb	112.6 kb	4.6 kb
	Others	9.9 kb	11.2 kb	48.3 kb	152.6 kb	
MUP						
	Haloferacales	0.8 kb	0.3 kb	0 kb	9.0 kb	1.1 kb
	Halobacteriales	1.9 kb	3.1 kb	6.3 kb	35.6 kb	2.4 kb
	Bacteroidetes+	0.3 kb	2.1 kb	1.9 kb	12.7 kb	1.1 kb
	Others	7.0 kb	8.5 kb	20.8 kb	72.7 kb	

Note: "Bacteroidetes+" contain identified and/or unidentified orders of Bacteroidetes

Appendix 10: Comparative assessment of existing carbon fixation enzymes of rTCA cycle in studied sites of Danakil Depression

ko00710: carbon fixation in photosynthetic organisms

KO	Enzyme description	EC number	DAL	GAL	BLA	LADP5	LADP6	LA	MUP
K00024	Malate dehydrogenase (MDH)	EC:1.1.1.37	√	√	√	√	√	√	√
K00134	Glyceraldehyde 3-phosphate dehydrogenase (GAPDH)	EC:1.2.1.12	√	X	√	√	√	√	√
K00615	Transketolase (TK)	EC:2.2.1.1	√	X	√	√	√	√	√
K00927	Phosphoglycerate kinase (PGK)	EC:2.7.2.3	√	X	√	√	√	√	√
K01006	Pyruvate,orthophosphate dikinase (ppdK)	EC:2.7.9.1	√	X	√	√	√	√	√
K01624	Fructose-bisphosphate aldolase, class II (FBA)	EC:4.1.2.13	√	X	√	√*	√*	√*	√*
K01783	Ribulose-phosphate 3-epimerase (rpe)	EC:5.1.3.1	√	X	√	√	√	√	X
K01803	Triosephosphate isomerase (TIM)	EC:5.3.1.1	√	X	√	√	√	√	√
K01808	Ribose 5-phosphate isomerase B (rpiB)	EC:5.3.1.6	√**	X	√	X	√**	X	X
K02446	Fructose-1,6-bisphosphatase II (glpX)	EC:3.1.3.11	√	X	√	√***	√	√***	√***
K01601	Ribulose bisphosphate carboxylase large chain	EC 4.1.1.39	√	√	√	√	√	X	X
K01602	Ribulose bisphosphate carboxylase small chain	EC 4.1.1.39	X	√	√	X	X	X	X
K01672	Carbonic anhydrase	EC 4.2.1.1	√	√	√	X	√	X	X

√* Only Fructose-bisphosphate aldolase, class I was detected

√** Only Ribose 5-phosphate isomerase A(rpiA) was detected

√*** Only Fructose-1,6-bisphosphatase, type I (EC 3.1.3.11) was detected

ko00720: carbon fixation pathways in prokaryotes

KO	Enzyme description	EC number	DAL	GAL	BLA	LADP5	LADP6	LA	MUP
K00024	Malate dehydrogenase (MDH)	EC:1.1.1.37	√	√	√	√	√	√	√
K00031	Isocitrate dehydrogenase (IDH1)	EC:1.1.1.42	√	X	√	√	√	√	√
K00031	Isocitrate dehydrogenase (IDH2)	EC:1.1.1.42	X	X	X	X	X	X	X
K00174	2-oxoglutarate ferredoxin oxidoreductase subunit alpha (korA)	EC:1.2.7.3	X	X	X	X	X	X	X
K00175	2-oxoglutarate ferredoxin oxidoreductase subunit beta (korB)	EC:1.2.7.3	X	X	X	X	X	X	X
K00239	Succinate dehydrogenase flavoprotein subunit (sdhA)	EC:1.3.99.1	√	X	√	√	√	√	√
K00240	Succinate dehydrogenase iron-sulfur subunit (sdhB)	EC:1.3.99.1	√	X	√	√	√	√	√
K00241	Succinate dehydrogenase cytochrome b556 subunit (sdhC)	EC:1.3.99.1	X	X	X	X	X	X	X
K00242	Succinate dehydrogenase membrane anchor subunit (sdhD)	EC:1.3.99.1	X	X	X	X	X	X	X
K00244	Fumarate reductase flavoprotein subunit (frdA)	EC:1.3.99.1	X	X	X	X	X	X	X
K00245	Fumarate reductase flavoprotein subunit (frdB)	EC:1.3.99.1	X	X	X	X	X	X	X
K00246	Fumarate reductase subunit C (frdC)	EC:1.3.99.1	X	X	X	X	X	X	X

K00247	Fumarate reductase subunit D (frdD)	EC:1.3.99.1	X	X	X	X	X	X	X
K00626	Acetyl-CoA C-acetyltransferase (2.AA440)	EC:2.3.1.9	√	√	√	X	√	√	√
K00925	Acetate kinase (ackA)	EC:2.7.2.1	√	√	√	X	√	√	X
K01006	Pyruvate,orthophosphate dikinase (ppdK)	EC:2.7.9.1	√	X	√	√	√	√	√
K01491	Methylenetetrahydrofolate dehydrogenase (NADP+)//Methenyltetrahydrofolate cyclohydrolase (fold)	EC:1.5.1.5 EC:3.5.4.9	√	X	√	√	√	√	√
K01679	Fumarate hydratase, class II	EC:4.2.1.2	√	X	√	√	√	√	√
K01681	Aaconitate hydratase (ACO)	EC:4.2.1.3	√	X	√	√	√	√	√
K01782	3-hydroxyacyl-CoA dehydrogenase//enoyl-CoA hydratase//3-hydroxybutyryl-CoA epimerase (fadJ)	EC:1.1.1.35	√	√	√	√	√	√	√
K01847	Methylmalonyl-CoA mutase (MUTB) AA750	EC:5.4.99.2	√	X	√	√	√	√	√
K01895	Acetyl-CoA synthetase (ACSS)	EC:6.2.1.1							
K01902	Succinyl-CoA synthetase alpha subunit (sucD)	EC:6.2.1.5	√	X	√	X	√	√	√
K01903	Succinyl-CoA synthetase beta subunit (sucC)	EC:6.2.1.5	√	X	√	√	√	√	√
K01958	Pyruvate carboxylase (PC)	EC:6.4.1.1	√	X	√	√	√	√	√
K01963	Acetyl-CoA carboxylase carboxyl transferase subunit beta (accD)	EC:6.4.1.2	X	X	X	X	X	X	X
K01965	Propionyl-CoA carboxylase alpha chain (PCCA)	EC:6.4.1.3	X	X	X	X	X	X	X
K01966	Propionyl-CoA carboxylase beta chain (PCCB)	EC:6.4.1.3	√	X	√	√	√	√	√
K03518	Carbon-monoxide dehydrogenase small subunit	EC:1.2.99.2	X	X	X	X	X	X	X
K03519	Carbon-monoxide dehydrogenase medium subunit	EC:1.2.99.2	X	X	X	X	X	X	X
K03520	Carbon-monoxide dehydrogenase large subunit	EC:1.2.99.2	X	X	X	X	X	X	X
K05606	Methylmalonyl-CoA/ethylmalonyl-CoA epimerase (MCEE)	EC:5.1.99.1	X	X	X	X	X	X	X
K13788	Phosphate acetyltransferase (PTA)	EC:2.3.1.8	√	√	√	√	X	X	X
K01644	Citrate lyase beta chain	EC 4.1.3.6	√	X	√	√	X	X	√
K00170	Pyruvate:ferredoxin oxidoreductase, beta subunit	EC 1.2.7.1	X	X	X	√	√	√	√
K00171	Pyruvate:ferredoxin oxidoreductase, delta subunit	EC 1.2.7.1	X	X	X	√	√	√	√
	α-ketoglutarate:ferredoxin oxidoreductase		X	X	X	X	X	X	X
K00240	Succinate dehydrogenase iron-sulfur protein/ fumarate reductase	EC 1.3.99.1	√	X	√	√	√	√	√
	2-oxoglutarate ferredoxin:oxidoreductase	EC 1.2.7.3	X	X	X	X	X	X	X

ko00680 : Methane metabolism

K03390	CoB--CoM heterodisulfide reductase subunit C	EC 1.8.98.1	X	√	X	X	X	X	X
K16179	Dimethylamine methyltransferase corrinoid protein		X	√	X	X	X	X	X
K16178	Dimethylamine:corrinoid methyltransferase		X	√	X	X	X	X	X
K00319	F420-dependent methylenetetrahydromethanopterin dehydrogenase	EC 1.5.99.9	X	√	X	X	X	X	X
K11260	Formylmethanofuran dehydrogenase (molybdenum) operon gene G		X	√	X	X	X	X	X
K00202	Formylmethanofuran dehydrogenase (molybdenum) subunit C	EC 1.2.99.5	X	√	X	X	X	X	X
K00203	Formylmethanofuran dehydrogenase (tungsten) subunit D	EC 1.2.99.5	X	√	X	X	X	X	X
K00201	Formylmethanofuran dehydrogenase subunit B	EC 1.2.99.5	X	√	X	X	X	X	X
K00672	Formylmethanofuran--tetrahydromethanopterin N-formyltransferase	EC 2.3.1.101	X	√	X	X	X	X	X
K00402	Methyl coenzyme M reductase gamma subunit	EC 2.8.4.1	X	√	X	X	X	X	X
K00401	Methyl coenzyme M reductase I beta subunit	EC 2.8.4.1	X	√	X	X	X	X	X
K00402	Methyl coenzyme M reductase I gamma subunit	EC 2.8.4.1	X	√	X	X	X	X	X
K00399	Methyl coenzyme M reductase II alpha subunit	EC 2.8.4.1	X	√	X	X	X	X	X
K00402	Methyl coenzyme M reductase II gamma subunit	EC 2.8.4.1	X	√	X	X	X	X	X
K01968	Methylcrotonyl-CoA carboxylase biotin-containing subunit	EC 6.4.1.4	X	√	X	X	X	X	X
K20509	Methylmalonyl-CoA decarboxylase, beta chain	EC 4.1.1.41	X	√	X	X	X	X	X
K08963	Methylthioribose-1-phosphate isomerase	EC 5.3.1.23	X	√	X	X	X	X	X
K16176	Monomethylamine methyltransferase corrinoid protein		X	√	X	X	X	X	X
K16238	Monomethylamine permease		X	√	X	X	X	X	X
K16176	Monomethylamine:corrinoid methyltransferase		X	√	X	X	X	X	X
	pyrrolysine-containing		X	√	X	X	X	X	X
K00583	N5-methyltetrahydromethanopterin:coenzyme M methyltransferase subunit G	EC 2.1.1.86	X	√	X	X	X	X	X
K00584	N5-methyltetrahydromethanopterin:coenzyme M methyltransferase subunit H	EC 2.1.1.86	X	√	X	X	X	X	X
K01499	N(5),N(10)-methenyltetrahydromethanopterin cyclohydrolase	EC 3.5.4.27	X	X	X	X	√	X	X

Appendix 11: ORFs with more than 50% coverage for protein families in Methanogenesis and Methanogenesis from methylated compound subsystems

EC number	Enzyme description	Taxonomic representation (Family level)	Coverage (%)	Identity (%)
EC 1.8.98.1	CoB--CoM heterodisulfide reductase subunit C	Burkholderiaceae	71	100
	Dimethylamine methyltransferase corrinoid protein	Bradyrhizobiaceae	100	95
	Dimethylamine:corrinoid methyltransferase	Bradyrhizobiaceae	100	91
EC 1.5.99.9	F420-dependent methylenetetrahydromethanopterin dehydrogenase	Burkholderiaceae*	18	100
	Formylmethanofuran dehydrogenase (molybdenum) operon gene G	Bradyrhizobiaceae	95	52
EC 1.2.99.5	Formylmethanofuran dehydrogenase (molybdenum) subunit C	Bradyrhizobiaceae	81	86
EC 1.2.99.5	Formylmethanofuran dehydrogenase (tungsten) subunit D	Enterobacteriaceae	50	100
EC 1.2.99.5	Formylmethanofuran dehydrogenase subunit B	Comamonadaceae, Bradyrhizobiaceae	59, 99.5	99, 91
EC 2.3.1.101	Formylmethanofuran--tetrahydromethanopterin N-formyltransferase	Bacteria	100	100
EC 2.8.4.1	Methyl coenzyme M reductase gamma subunit	Comamonadaceae, Bradyrhizobiaceae	100, 63	100, 91
EC 2.8.4.1	Methyl coenzyme M reductase I beta subunit	Burkholderiaceae	90	81
EC 2.8.4.1	Methyl coenzyme M reductase I alpha subunit	Bradyrhizobiaceae	100	83-96
EC 2.8.4.1	Methyl coenzyme M reductase II alpha subunit	Thermaceae*	26	97
EC 2.8.4.1	Methyl coenzyme M reductase I gamma subunit	Burkholderiaceae	73	98
EC 2.8.4.1	Methyl coenzyme M reductase II gamma subunit	Burkholderiaceae*	41	98
	Monomethylamine methyltransferase corrinoid protein	Bradyrhizobiaceae	97	80
	Monomethylamine permease	Bacilli	100	100
	Monomethylamine:corrinoid methyltransferase	Burkholderiaceae	82	96
	pyrrolysine-containing	Comamonadaceae	98	79.5
EC 2.1.1.86	N5-methyltetrahydromethanopterin:coenzyme M methyltransferase subunit G	Propionibacteriaceae, Bradyrhizobiaceae	100, 68	99, 79

EC 2.1.1.86	N5-methyltetrahydromethanopterin:coenzyme M methyltransferase subunit H	Rhodothermus	87	76

* No alternative OTU was detected

Appendix 12: Optical density measurements of Acidic and Basic culture media

OD (380nm) measurement of replicated basic culture media where sterile control was used to blank the UV-VIS spectrometer

	Replicate I	Replicate II	Replicate III
0hr	0.003	0.005	0.002
24hr	0.119	0.087	0.099
48hr	0.308	0.206	0.284
72hr*	-0.173	-0.468	-0.35
96hr*	-0.107	-0.275	-0.174
120hr*	-0.291	-0.3	-0.21

* the control reference used as a blank is remains thicker thus absorb more than the sample which utilizes the substrates for growth and resulted in negative values

OD (380nm) measurement of replicated basic culture media and abiotic control where sterile media without FeCl3 used to blank the UV-VIS spectrometer

	Control	Replicate I	Replicate II	Replicate III
0hr	1.103	1.114	1.107	1.111
24hr	1.271	1.403	1.373	1.388
48hr	1.49	1.806	1.704	1.758
72hr	2.046	1.952	1.59	1.603
96hr	1.797	1.72	1.532	1.6
120hr	1.822	1.604	1.481	1.58

OD (600nm) measurement of replicated acidic culture media where sterile control was used to blank the UV-VIS spectrometer

	Replicate I	Replicate II	Replicate III
0hr	0.001	0.009	0.005
24hr	0.418	0.245	0.331
48hr	0.493	0.369	0.431
72hr	0.538	0.338	0.452
96hr	0.755	0.534	0.664
120hr	0.53	0.473	0.582