ABSTRACT

BOZDAG, AHMET. Investigation of Methanol and Formaldehyde Metabolism in *Bacillus methanolicus*.(Under the direction of Prof. Michael C. Flickinger).

Bacillus methanolicus is a Gram-positive aerobic methylotroph growing optimally at 50-53 °C. Wild-type strains of *B. methanolicus* have been reported to secrete 58 g/l of L-glutamate in fed-batch cultures. Mutants of *B. methanolicus* created via classical mutangenesis can secrete 37 g/l of L-lysine, at 50 °C. The genes required for methylotrophyin *B. methanolicus* are encoded by an endogenous plasmid, pBM19 in strain MGA3, except for hexulose phosphate synthase (*hps*) and phosphohexuloisomerase (*phi*) which are encoded on the chromosome.It is a promising candidate for industrial production of chemical intermediates or amino acids from methanol. *B. methanolicus* employs the ribulose monophospate (RuMP) pathway to assimilate the carbon derived from the methanol, but enzymes that dissimilate carbon are not identified, although formaldehyde and formate were identified as intermediates by ^{13}C NMR. It is important to understand how methanol is oxidized to formaldehyde and then, to formate and carbon dioxide. This study aims to elucidate the methanol dissimilation pathway of *B. methanolicus*.

Growth rates of *B. methanolicus* MGA3 were assessed on methanol, mannitol, and glucose as a substrate. *B. methanolicus* achieved maximum growth rate, μ_{max}, when growing on 25 mM methanol, 0.65±0.007 h^{-1}, and it gradually decreased to 0.231±0.004 h^{-1} at 2Mmethanol concentration which demonstrates substrate inhibition. The maximum growth rates (μ_{max}) of *B. methanolicus* MGA3 on mannitol and glucose are 0.532±0.002 and

0.336±0.003 h^{-1}, respectively. Spiking with 100 mM methanol did not cause any growth perturbation of*B. methanolicus* MGA3 when grown on methanol, whereas, 50 mM methanol was enough to halt the growth when it was grown on mannitol. However, growth of *B. methanolicus* MGA3 on methanol was halted completely by a 1 mM formaldehyde spike, while mannitol grown *B. methanolicus* MGA3 can tolerate 2 mM formaldehyde addition. Formate did not have any apparent growth perturbations at tested levels (up to 2 mM).

B. methanolicus MGA3 has a small genome, ~3.3 Mb, and contains 2,773 ORFs larger than 100 bp. Its genome encodes two methanol dehydrogenase(*mdh*) genes beside the *mdh* on pBM19. All of the RuMP pathway genes encoded on the pBM19 have counterparts encoded on the chromosome also. The genome contains putative tetrahydrofolate (THF)-linked enzymes for formaldehyde oxidation besides putative aldehyde dehydrogenase(*aldh*) genes. Two putative formate dehydrogenase (*fdh*) genes were also identified in the genome as were the genes of the Wood-Ljungdahl pathway genes.*B. methanolicus* strain PB1 was also found to carry a plasmid which we named as pBM20. The difference in the formaldehyde dissimilation pathway of strain PB1 was shown via Southern Blot. The PB1 strain also has 2-fold higher formaldehyde dehydrogenase and 3-fold higher formate dehydrogenase activity compared to strain MGA3.

The methanol dissimilation pathway genes,- putative THF-linked formaldehyde oxidation genes, putative *aldh* genes, putative *fdh* gene- are all down regulated in response to spikes of methanol, formaldehyde and formate. However, the regulation of RuMP pathway genes were different in methanol vs. mannitol grown *B. methanolicus* MGA3 in that they were down regulated from already high levels in methanol grown cells whereas they were upregulated in mannitol grown cells. Formate is a more potent regulatory compound compared to methanol and formaldehyde.

Investigation of Methanol and Formaldehyde Metabolism in *Bacillus Methanolicus*

by
Ahmet Bozdag

A dissertation submitted to the Graduate Faculty of
North CarolinaStateUniversity
in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Microbiology

Raleigh, North Carolina

2013

APPROVED BY:

Dr. Michael C. Flickinger Committee Chair	Dr. Michael R. Hyman
Dr. Jose Manuel Bruno-Barcena	Dr. Amy M. Grunden

DEDICATION

To my wife Turkan and daughter Vildan

BIOGRAPHY

Ahmet Bozdag was born in Besni, Southeast Turkey as the youngest of a large family. He moved to Gaziantep after elementary school, to seek a better education at the age of 11. He completed his high school and attended Middle East Technical University (METU) in Ankara. He taught science in middle and high schools in Adiyaman and Ankara while seeking a graduate degree at METU. Winning the Diversity Visa lottery, or Green Card, he moved to United States with his ambition to seek a doctoral degree in microbiology. He earned his Master's Degree at Georgia College & State University before attending the Microbiology department at NCSU during which he married to a beautiful woman, Turkan. He had his first child, a gorgeous baby girl Vildan just before completing his PhD degree.

ACKNOWLEDGMENTS

I would like to thank professor Michael C. Flickinger foremost, my thesis advisor, for accepting me into his laboratory and for his invaluable guidance. I would also like to thank my thesis committee members Drs. Jose Manuel Bruno-Barcena, Amy M. Grunden, Michael B. Goshe, and Michael R. Hyman.

I would also like to thank my colleagues in the Flickinger laboratory, past and present. I would like to thank Dr. Thomas Pribyl for his partnership in the project, Dr. Jimmy Gosse for his ideas and discussions, Dr. Jessica Jenkins, Oscar I. Bernal and Mark Schulte for their friendship and ideas. I would like to thank German visiting students who helped me on the project along the way; Paul Simon, Fabian Gutmacher, Felix Hart and Elisa Degenkolbe. I would like to thank specially Dr. Claire Komives of San Jose State University for her help with sequencing the*B. methanolicus* genome.

Also, I would like to thank Microbiology department staff members; Cindy Whitehead for reminding me everything I forgot and Robert Davis for his friendship.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

Methylotrophic microorganisms

Methylotrophs are those microorganisms able to grow at the expense of reduced carbon compounds containing one or more carbon atoms but no carbon-carbon bonds (Anthony, 1982). This is a broad definition and includes bacteria capable of growth on methane, methanol, methylated amines, halogenated methanes and methylated sulfur compounds. Obligate methylotrophs exclusively utilize C_1 compounds as a sole carbon and energy source while facultative methylotrophs can utilize multicarbon compounds as well. Although still considered to be methylotrophic bacteria, the group that is able to catalyze the conversion of methane to methanol is further characterized as methanotrophic bacteria. Although methylotrophic microorganisms were known to be Gram-negative, obligate aerobes until 1983 (Haber et. al., 1983), this statement has been broadened by the characterization of Gram-positive methylotrophs.

Habitats of methylotrophs

Methylotrophic bacteria are ubiquitous in the environment and have a significant contribution to the global carbon cycle. They can be isolated from a wide range of habitats. Nearly all samples taken from mud, swamps, rivers, rice paddies, oceans, and ponds, soils from meadows, and streams, sewage sludge, and several other habitats yielded methylotrophs in enrichments (Hanson, 1992). Although their presence may be attributed to the abundance of methane in these environments because of methanogens, methanol produced as the end product of pectin metabolism also offers a suitable niche for methylotrophs to thrive. Methanol is the major end product of pectin metabolism by pectinolytic microorganisms (Schink and Zeikus, 1980), although they do not consume it. Methanol concentrations are stoichiometrically coupled to pectin metabolism in pectinolytic microorganisms whether it is aerobic or anaerobic (Schink and Zeikus, 1980). Another compound that may serve as a source of methylotrophy is lignin. Formaldehyde, but not methanol, is the byproduct of lignin biodegradation (Kuwahara, H., 1980). Most methylotrophs, excluding organisms with RuBP, assimilate carbon at the level of formaldehyde. Methanol can be used as carbon, hydrogen and oxygen sources by these microorganisms.

Different techniques have been employed to detect methylotrophic microbial communities in various environments. Use of stable-isotope probing (SIP) demonstrated the presence of *Methylobacterium*, *Methylophaga* and *Bacillus spp*. in the methane, methanol and methylamine microcosms respectively, in Lonar Lake, which is a unique saline and alkaline ecosystem formed by a meteor impact some 52,000 years ago (Antony et. al., 2010). The same technique also detected a broad range of methylotrophs in samples from Lake Washington (Nercessian et. al., 2005 and Kalyuzhnaya et al., 2008), acidic forest soil (Radajewski et. al., 2002) marine samples (Neufeld et. al., 2008) and an alkaliphilic saline soda lake (Doronina et al., 2003). All these studies demonstrate the broad habitat of methylotrophs from fresh water lake sediments (Nercessian et. al., 2005) to soils and sediments from Antarctica (Azra et. al., 2005). Another habitat that was not thought to be associated with methylotrophic microorganisms, the human mouth, was also shown to contain methylotrophs. Methylotrophs from the human mouth were detected using 16S rRNA gene sequencing and primers designed for the *mxaF* gene of methanol dehydrogenase (Anesti et. al., 2005). The identification and isolation of methylotrophic microorganisms is not complete as new microorganisms are still being isolated (Boden et. al., 2008).

Carbon assimilation in methylotrophic microorganisms

The first reaction in methylotrophy is the oxidation of methanol to formaldehyde. Formaldehyde is formed from methanol via oxidation by methanol dehydrogenase or methanol oxidase. From this oxidation level most of the carbon is assimilated into cell biomass. There are four different pathways by which methylotrophic organisms can assimilate carbon substrates into cell material (Figure 1.1). One of the routes, the serine cycle, involves carboxylic acids and amino acids as intermediates. The other three routes, ribulosebisphosphate (RuBP), known as the Calvin cycle, ribulose monophosphate (RuMP), and xylulose monophosphate- also known as dihydroxyacetone pathway (XuMP), involve carbohydrate intermediates. The XuMP is found only in yeast grown on methanol. The C_1 carbon can be assimilated as different intermediates. RuMP and XuMP pathways incorporate carbon at the level of formaldehyde while the serine cycle can assimilate both formaldehyde and carbon dioxide. The distribution of assimilation pathways in methylotrophs is shown in Figure 1.2. In RuBP organisms the source of fixed carbon is the carbon dioxide. The methylotrophic bacteria are grouped according to the way they assimilate formaldehyde. The Type I microorganisms employ the ribulose monophosphate (RuMP) pathway whereas Type II microorganisms use the serine pathway for assimilation of formaldehyde. Some microorganisms like *M.capsulatus* (Bath), are classified as Type X since they also contain some enzymes of the serine cycle while carbon is assimilated via the RuMP pathway (Hanson and Hanson, 1996). The energy required for the assimilation reactions is obtained by oxidation of a portion of the formaldehyde to carbon dioxide which produces NADH.

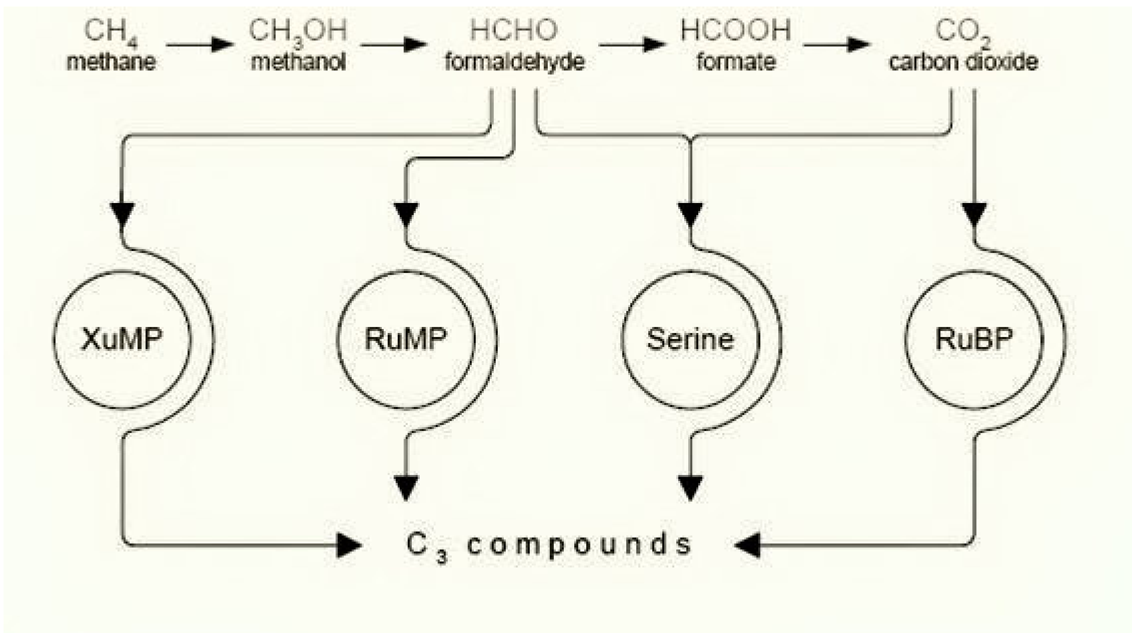


Figure 1.1. Four different pathways for one-carbon assimilation in methylotrophs (Dijkhuizen*et. al.*,1992).

Ribulosebiphosphate pathway, RuBP

The ribulosebisphosphate (RuBP) pathway, also known as the Calvin cycle, is the least common route for assimilation of carbon in methylotrophs. The carbon is assimilated at the level of carbon dioxide and is mainly employed by chemolithotrophic bacteria. Carbon dioxide is condensed to ribulose 1,5-bisphosphate to produce two molecules of 3-phosphoglycerate by the key enzyme of the cycle, ribulosebisphosphate carboxylase. This molecule is then reduced to glyceraldehyde-3-phosphate. One molecule of glyceraldehyde-3-phosphate is isomerized to dihydroxyacetone phosphate and then condensed with another glyceraldehyde-3-phosphate to yield a fructose-1,6-diphosphate (F6P). One molecule of F6P is gained from 6 molecules of carbon dioxide consuming 18 ATP and 12 NAD(P)H.

Table 1.1**.**Distribution of assimilation pathways in methylotrophs (Anthony, 1991).

Ribulose biphosphate (RuBP) Pathway

- *Paracoccus denitrificans*
- *Thiobacillus* sp.
- *Rhodopseudomonas acidophila*
- *Xanthobacter* sp.
- *Rhodobacter* sp.

Xylulose monophosphate (XuMP) Pathway

- *Hansenula* sp.
- *Candida* sp.
- *Pichia* sp.
- *Torulopsis* sp.

Ribulose monophoshpate (RuMP) Pathway

- *Methylococcus capsulatus*
- *Methylophilus methylotrophus*
- *Arthrobacter* P1
- *Bacillus* sp.
- *Acetobacter methanolicus* MB58

Serine Pathway

- *Methylosinus* sp.
- *Methylobacterium extorquens* AM1
- *Hyphomicrobium*

Xylulose monophosphate pathway, XuMP

The xylulose monophosphate cycle is a yeast specific pathway in which the net conversion of three formaldehyde molecules into a C_3 compound, dihydroxyacetone phosphate, is accomplished. Yeast was the first eukaryotic organism shown to grow on C_1 compounds in 1969 (Ogata et. al., 1969). The methylotrophic yeasts isolated from natural sources appear to belong to a limited number of genera. Screening of type-culture collections conducted by different research groups showed that only 15 species within the genera of *Candida*, *Hansenula*, *Pichia* and *Torulopsis* were capable of growth on methanol (Hazeu et. al., 1972, and Oki et. al., 1972). Lee and Komagata conducted a taxonomic study of 31 fresh isolates and 22 strains known to assimilate methanol (Lee and Komagata, 1980). Besides confirming the above results, they also observed that the majority of the strains examined were able to assimilate pectin, which yields methanol upon hydrolysis (Lee and Komagata, 1980).

The conversion of methanol into formaldehyde is catalyzed by alcohol oxidases and also yields hydrogen peroxide as a byproduct. While this reaction takes place in the peroxisomes, enzymes for the dissimilation of formaldehyde are located in the cytosol. The formaldehyde diffuses into the cytosol and condenses with reduced glutathione (GSH) non-enzymatically, to form S-hydroxymethylglutathione. The product S-glutathione is either hydrolyzed into formate and oxidized to carbon dioxide by formate dehydrogenase (Allais and Louktibi, 1983 and, Schütte et. al., 1976) or oxidized directly. Another enzyme, formaldehyde reductase, which reduces formaldehyde to methanol, has been detected in methylotrophic yeast which is thought to serve as a “safety valve” when the cytosolic formaldehyde concentration is too high (Trotsenko et. al., 1984).

Assimilation of formaldehyde is accomplished by the xylulose monophosphate (XuMP) cycle. Dihydroxyacetone (DHA) synthase was shown to be the first enzyme in the assimilation process which takes place in the peroxisome. This enzyme is a special transketolase that catalyzes the glycol-aldehyde transfer from xylulose-5-phosphate as a donor to the acceptor molecule formaldehyde in a reversible reaction (O'Connor and Quayle, 1980).

Ribulose monophosphate pathway, RuMP

In the ribulose monophosphate (RuMP) pathway, the carbon from methane or methanol is assimilated at the level of formaldehyde. In this pathway, one molecule of pyruvate or dihydroxyacetone is synthesized from three molecules of formaldehyde. The pathway can be divided into 3 parts; fixation, cleavage and rearrangement. In the first or fixation part, the formaldehyde is condensed with ribulose-5-phosphate to give hexulose phosphate which is then isomerized to fructose-6-phosphate. These steps are catalyzed by hexulose phosphate synthase (HPS) and hexulose phosphate isomerase (HPI), respectively. In the second part of the pathway, one molecule of F6P is converted either to 2-keto-3-deoxy 6-phosphogluconate (KDPG) or to fructose biphosphate (FBP). These molecules are then cleaved by aldolases to glyceraldehyde 3-phosphate plus either pyruvate (KDPG) or dihydroxyacetone (FBP). In the rearrangement part of the pathway, the glyceraldehyde phosphate and F6P undergo a series of reactions which lead to regeneration of ribulose 5-phosphate. There are two variants in this step; the sedoheptulosebiphosphate variant and the transalddolase variant. The most common variant is the KDPG aldolase/transaldolase (TA) and less common variant is the FBP aldolase/SBP phosphatase. The energetics are shown in the Table 1.2.

Table 1.2. Energetics of the RuMP pathway of formaldehyde fixation. Pyruvate is considered the end-product (Dijkhuizen *et. al.*,1992)

	Energy change (3HCHO $\longrightarrow$ 1 pyruvate)	
Variant	$NAD(P)H_2$	ATP
KDPGA/TA	+1	0
FBPA/SBPase	+1	0
KDPGA/SBPase	+1	-1
FBPA/TA	+1	+1

Serine cycle

Studies with short term incubations of *Pseudomonas* AM1, which was renamed*Methylobacterium extorquens* AM1, with ^{14}C-labeled methanol resulted in sequential labeling of serine followed by glycine, malate and aspartate leading to the hypothesis that the serine cycle was a C_1 assimilatory pathway (Large et. al., 1961). Unlike the RuMP pathway in which carbohydrates serve as intermediates, the serine cycle utilizes carboxylic acids and amino acids. Carbon is assimilated via formaldehyde and carbon dioxide. The serine cycle genes in *M. extorquens* (AM1) are transcribed from three separate units that are positively regulated by a LysR-Type transcriptional regulator that is homologous to CbbR (Kalyuzhnaya andLidstrom. 2003). Glycine is the starting compound in the serine cycle, which is condensed with formaldehyde by serine trans-hydroxymethylase to yield serine. Serine-glycine aminotransferase converts serine to hydroxypyruvate which is then reduced to glycerate at the expense of NADH. Phosphoglycerate is one of the compounds from which carbon can be drawn for cell material and it is produced from glycerate at the expense of 1 ATP, catalyzed by glycerate kinase. Phosphoenolpyruvate (PEP) is the compound that condenses with the carbon dioxide to form oxaloacetate and then malate. Malate then condenses with CoA giving rise to malyl-CoA. Malyl-CoA is then split into glyoxylate and acetyl-CoA by malyl-CoA lyase. Glyoxylate is finally converted to glycine by serine-glycine aminotransferase which completes the cycle. In isocitratelyase positive (*icl*$^{+}$) organisms, the acetyl-CoA is oxidized to glyoxylate with a set of reactions that involves isocitratelyase and this pathway also occurs in bacteria growing on ethanol or acetate. However, most of the methylotrophic microorganisms are known to be *icl*$^{-}$.

Regeneration of glyoxylate from acetyl-CoA in *icl*$^{-}$ microorganisms remained an enigma until a new pathway was recently proposed. Erb and coworkers have shown the transformation of isotopically labeled acetate and bicarbonate to ethylmalonyl-CoA by cell extracts from cultures grown in acetate by the isocitratelyase negative bacterium *Rhodobactersphaeroides* (Erb et. al., 2007). They identified crotonyl-CoA carboxylase/reductase as a key enzyme in the ethylmalonyl-CoA pathway which catalyzes the condensation of carbon dioxide with crotonyl-CoA to yield ethylmalonyl-CoA. In short (Figure 1.2), two acetyl-CoA molecules and 1 molecule of carbon dioxide condense yielding a C5 compound methylmalyl-CoA which is then split into propionyl-CoA (C3) and glyoxylate (C2). Glyoxylate enters into the serine cycle while propionyl-CoA is carboxylated to a C4 compound eventually yielding succinate which then enters the serine cycle as malate.

Putative genes encoding for these enzymes were also identified in the *M.extorquens* AM1 genome (Vuilleumier et. al., 2009), but not in the Type X methylotroph *M. capsulatus* (Bath) genome (Ward et. al., 2004).

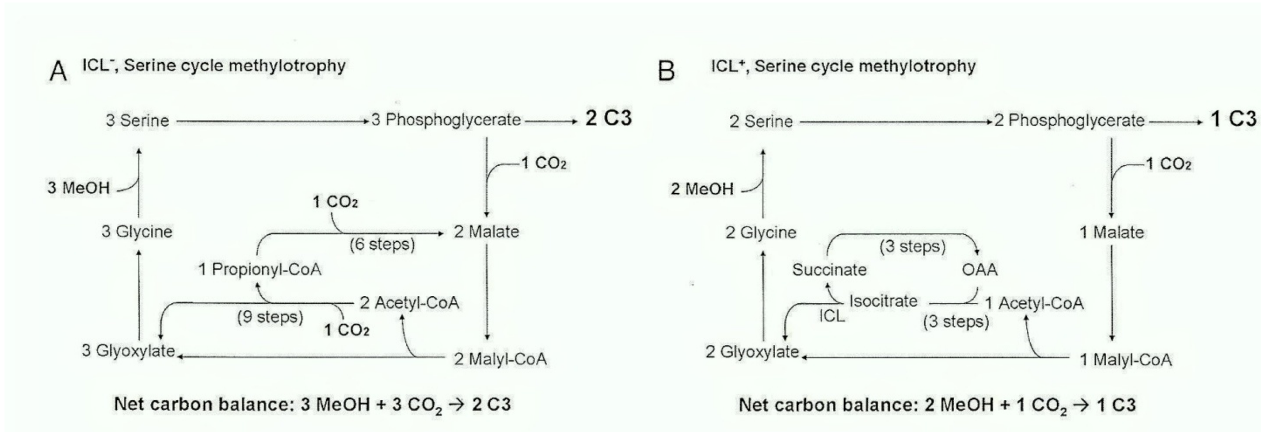


Figure 1.2. Comparison of C1 assimilation pathways by isocitratelyase-negative (A) and positive (B) methylotrophs (Peyraud et. al., 2009).

Dissimilation of carbon in methylotrophs

Methanol dehydrogenase

The first step in the methylotrophy is the oxidation of methanol into formaldehyde. This conversion is carried out by methanol dehydrogenase in bacteria and by methanol oxidase in yeast. Only the methanol dehydrogenase will be summarized here. Cell lysis achieved by using the French press or by sonication provides sufficient enzyme for purification, but when assayed, some MDH is also found associated with the membranes. MDH has been found in membranes or in membrane vesicles in a wide variety of bacteria including *Methylococcus capsulatus* (Wadzinski and Ribbons, 1975), *Hyphomicrobium* X (Duine et al., 1978), and*Methylophilusmethylotrophus* (Cross and Anthony, 1980).

Pyrroloquinoline quinone (PQQ) is ubiquitous as a cofactor of alcohol dehydrogenases. Gram-negative microorganisms employ PQQ-containing methanol dehydrogenases like MDH of *Methylobacteriumextorquens* and *Methylophilus* sp. (Anthony and Williams, 2003). It has a $\alpha_2\beta_2$tetrameric structure and each α subunit contains one PQQ molecule and one Ca^{2+} ion. The αsubunit is around 66 kDA while the β subunit is 8.5 kDA (Anthony and Williams, 2003). The first step is the transfer of an electron to cyctochrome c_L through PQQ. The PQQ of MDH is reduced by methanol, and formaldehyde is released. Then, two electrons are transferred to cyctochrome c_L from $PQQH_2$ which oxidizes $PQQH_2$ back to quinone by the way of a free radical semiquinone (Anthony and Williams, 2003). Different mechanisms have been proposed as to how methanol is oxidized by MDH (Leopoldini et. al., 2007), but the overall reaction is thought to be exothermic (Zhang et. al., 2007).

Most emphasis has been put on the MDHs of Gram-negative bacteria since most of the research has been done on them, and there is less data regarding MDHs of the Gram-positive methylotrophs. Gram-positive bacteria lack a clear periplasmic space, and they do not posses PQQ (Hektor et. al., 2000). The methanol dehydrogenases of Gram-positive bacteria are different than those of Gram-negative in that they are NAD-dependent. They may constitute up to 22% of total protein when grown on methanol (Arfman et.al., 1992). The MDH of thermotolerant *Bacillus methanolicus* C1 is a decameric protein with an estimated size of 43 kDA (Vonck et. al., 1991). Each subunit contains one zinc and one or two magnesium ions. Each subunit also contains one tightly bound NAD(H) molecule as a cofactor. This NAD molecule is oxidized by formaldehyde and reduced by methanol while remaining bound to the protein (Arfman et al., 1997). The *B. methanolicus* C1 MDH also requires exogenous NAD that re-oxidizes the cofactor NADH. A protein called activator protein (ACT) increases the affinity of MDH for methanol and may enhance the methanol turnover rate up to 40 fold (Arfman et. al., 1991). The exogenous NAD and Mg^{2+} are strictly required for the function of the activator enzyme (ACT).

Acetobacterpasteurianus SKU 1108 is found to possess both PQQ-containing alcohol dehydrogenase (ADH) and NAD-dependent ADH. A mutant strain deficient in PQQ-containing ADH has elevated levels of NAD-dependent ADH and grows better on ethanol which leads to the conclusion that PQQ-dependent ADHs function in catabolic acetate production while NAD-dependent ADHs function in assimilation of ethanol (Chinnawirotpisanet.al., 2003).

Formaldehyde

Formaldehyde is the simplest of the aldehydes and ranks as one of the largest volume commodity chemicals with production of 29 million tons/yr in 2010 which is expected to grow 5% per year from 2011-2015 (Merchant Research & Consulting ltd, 2012). It is one of the most important basic chemicals required for manufacture of industrial and consumer products. It is industrially produced from methanol by different synthesis routes with up to 97-98 % reported yield (Reuss et. al., 2005). The largest amount of formaldehyde is used in the manufacture of resins for the production of adhesives and impregnating resins which are employed for manufacturing particle boards, plywood and furniture. These resins are also used as auxiliaries in the textile, leather, rubber and cement industries. In agriculture, formaldehyde is used as a fumigant and as a preservative in animal feed (O'Brien et. al., 2005). The second largest use of formaldehyde is as an intermediate for synthesizing other chemical compounds. This constitutes about 40% of total formaldehyde consumption while direct use only constitutes a small fraction of total consumption, 1.5%. As a C_1 compound, formaldehyde is irreplaceable with any other compound (Reuss, 2005).

A trace amount of formaldehyde is present in the air as a result of photochemical oxidation of hydrocarbons. It is also emitted to air from plastics and resin glues (WHO, 2005). In rural air, terpenes and isopropene emitted by foliage also react with hydroxyl radicals to form formaldehyde (Atkinson 1990). Formaldehyde concentrations of up to 30µg/L have been found in ozonated drinking water (Krasner et. al., 1976). Formaldehyde is also present in rain water (1.38 mg/L) (Kitchens et. al., 1989) and fog water (6.8 mg/L) (Muir, 1991). It is even found in the Antarctic troposphere up to 0.7 ppbv, fluctuating seasonally (Riedel, 1999). The general route of exposure to formaldehyde is through inhalation (US NRC, 1980). The concentration of formaldehyde in food is found to be in the range of 3-23 mg/kg (NTP, 2010). Trezl et. al., (1997) found the level of formaldehyde to range from 3.6 mg/kg to 35 mg/kg in different plants most of which are used as human food.

Endogenous resources of formaldehyde include lipid peroxidation, carbohydrate auto-oxidation, carbohydrate metabolism, amine oxidase-catalyzed metabolic activation and cytochrome P-450 catalyzed metabolic activation (O'Brien et. al., 2005). One other important source of formaldehyde is methanol, generated by methanol dehydrogenases in methylotrophic microorganisms. It is a key metabolite in methylotrophy and its concentration should be controlled since it is also very reactive and can be toxic to microorganisms.

Formaldehyde toxicity

Formaldehyde is the most reactive aliphatic aldehyde molecule and takes part in many well-known chemical reactions (Walker, 1975). This reactivity is the major reason for its toxicity. It also has mutagenic effects. With the use of xanthine guanine phosphoribosyltransferase gene as a genetic target, 4mM formaldehyde was found to induce 41% large insertions, 18% large deletions and 41% point mutations in *E. coli*(Crosby et. al., 1988). Formaldehyde forms monohydroxymethyl adducts with four major ribonucleosides which occur at the exocyclic amino groups of deoxyguanosine, deoxycytidine, deoxyadenosine and at the endocyclicimino group of deoxtymidine. (McGhee and von Hippel, 1975a and 1975b). Incubation of single-stranded polyadenine with formaldehyde at 23 °C results in the bonding of the formaldehyde at a ratio of 1.00±0.05 per base pair. The initial denaturation of DNA by formaldehyde occurs in AT-rich regions in the interior of the DNA molecule. The denaturation occurs as follows: (i) a small sequence of DNA base pairs opens as a result of local thermal fluctuations or for transcription, (ii) the exocyclic amino group of an adenine residue is exposed, (iii) the adenine reacts with formaldehyde which either reforms in a hydroxymethylated AT base pair or remains as unbound to the complementary strand depending on the temperature and other environmental factors (McGhee and von Hippel, 1977a). Once bound to the DNA, formaldehyde lowers the T_m of the DNA sequence and the change in T_m increases with the AT content of the DNA (McGhee and von Hippel, 1977b).

Reactions of formaldehyde with amino acids, peptides and proteins were investigated using ^{13}C and ^{15}N NMR. The reactions of formaldehyde with proteins start with the formation of methylol adducts on amino groups. The major derivatives formed with single amino acids are monohydroxymethyl adducts, dihydroxymethyl adducts and cyclic products (Heck et. al., 1990). The methylol adducts of primary amino groups are partially dehydrated, yielding Schiff-bases, which can form cross-links with several amino acid residues and produce irreversible covalently cross-linked complexes between proteins and single stranded DNA (Bolt, 1987). Thus, organisms should keep metabolic formaldehyde levels low to avoid these undesired reactions.

Formaldehyde oxidation

Studies of methylotrophic microorganisms growing on methanol have shown the need for keeping the methanol concentration limited since specific growth rates declined after methanol concentration in the culture medium of *B. methanolicus* reached a critical level of 12g/L (Jakobsen et. al., 2006). In the same study, it was shown that formaldehyde is responsible for the decrease of the specific growth rate by adding extracellular formaldehyde (Pluschkell, 1998) but there are no published studies reporting that the formaldehyde produced as a metabolic intermediate inhibits microbial growth. An increase in the level of intracellular formaldehyde exerts stress on the microorganisms which may lead to induction of multiple stress response pathways including general stress response mechanisms (Nguyen et. al., 2009). The level of formaldehyde can be lowered by either assimilation by the RuMP or the serine cycle or dissimilation by oxidation to carbon dioxide. The same study has shown the specific induction of the ribulose monophosphate (RuMP) pathway in response to formaldehyde among other carbonyl electrophiles in non-methylotrophic *B. subtilis* where it is thought to serve as a detoxification pathway (Nguyen et. al., 2009). Accordingly, bacteria have capabilities to dissimilate excess formaldehyde as carbon dioxide. Dissimilatory pathways also generate energy required for growth on C_1 compounds in methylotrophic organisms although the reactions may differ as described below.

Linear formaldehyde oxidation pathways by aerobic methylotrophic organisms are summarized in Figure 1.3. The simplest pathway that converts formaldehyde to formate is catalyzed by the dye-linked formaldehyde dehydrogenase (DL-FALD) enzyme. This enzyme in *M. capsulatus* (Bath) is associated with membranes, and it has been shown to be the major formaldehyde oxidizing enzyme when *M. capsulatus* (Bath) cultures proliferate in medium supplemented with high copper concentration (Zahn et. al., 2001). The redox cofactor of the DL-FALD has been identified as the pyrroloquinoline quinone (PQQ) (Zahn et. al., 2001). However, another compound that can act as cofactor is mycothiol. In Gram-positive methylotrophs, NAD- and mycothiol-linked FALD have been identified as the major formaldehyde oxidation enzymes consisting of a single type of apo-protein (van Ophem et. al., 1992). The presence of bacillithiol as a thiol containing antioxidant in *Bacillus* species was demonstrated recently (Newton et. al., 2009). However, to date, no FALD was shown to be bacillithiol-dependent.

The third pathway to oxidize formaldehyde to formate in Gram-negative methylotrophs such as *Paracoccusdenitrificans* takes three steps, two enzymes and glutathione (GSH) as a cofactor of the FALD (Ras et. al., 1995 and Harms et. al., 1996). In this case, it requires two enzymes to oxidize formaldehyde to formate. Formaldehyde spontaneously condenses with GSH to form S-hydroxylmethyl GSH which is then converted to S-formyl GSH by NAD-GSH dependent FALD. The second enzyme, S-formyl GSH hydrolase catalyzes the conversion into formate while releasing GSH.

There are two routes for folate-linked formaldehyde oxidation; THF- or THMPT-linked pathways. In the former route, THF spontaneously condenses with formaldehyde to yield methenyl-THF. This compound is oxidized to formate and then carbon dioxide by the following enzymes: methylene THF dehydrogenase, methenyl THF cyclohydrolase, formyl THF synthetase and finally formate dehydrogenase. The THMPT-linked pathway was thought to be specific for archaea, and its presence in bacteria was considered unlikely. However, Chistoserdova et al., showed the presence of these enzymes in *M. extorquens* AM1 (Chistoserdov et. al., 1994). Indeed, it was shown to be more prevalent in bacteria than previously thought (Vorholt et al., 1999). It is also present in *M. capsulatus* (Bath), in which DL-FALD is thought to be the major dissimilatory enzyme (Ward et. al., 2004 and Zahn et. al., 2001). Unlike condensation of formaldehyde with THF, which occurs spontaneously, an enzyme was suggested to help condense formaldehyde with THMPT, named formaldehyde activating enzyme (Fae) (Vorholt et. al., 2000). An analog of Fae was also suggested that helps condense formaldehyde with GSH which was termed glutathione-dependent formaldehyde-activating enzyme (Gfa) (Goenrich et. al., 2002)

THF-linked enzymes formate-THF ligase, methenyl-THF cyclohydrolase and methylene-THF ligase also function in the anaerobic Wood-Ljungdahl pathway where carbon dioxide and carbon monoxide are converted into acetyl-CoA. Acetogenic bacteria assimilate carbon dioxide into biomass via this route (Ragsdale, 2008).

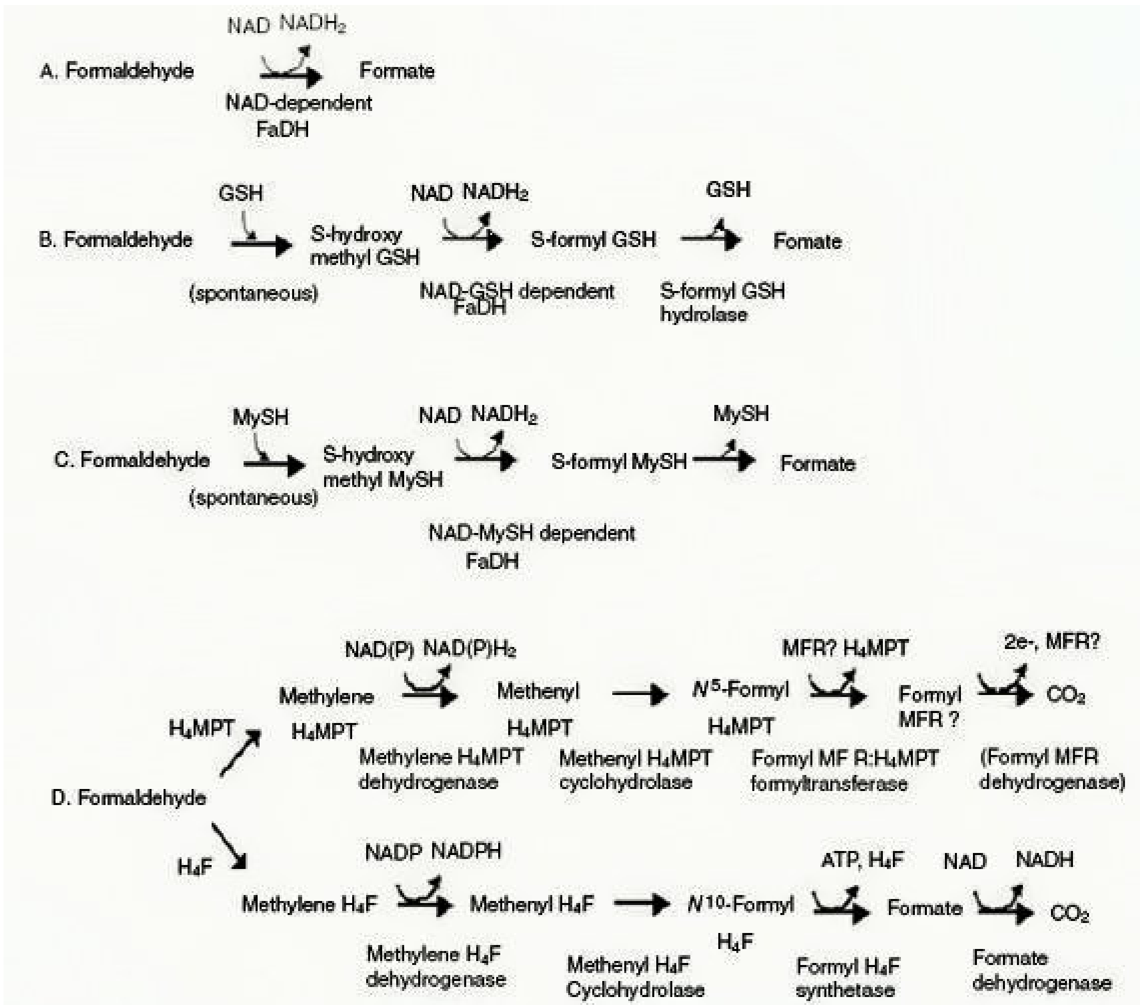


Figure 1.3. Formaldehyde oxidation pathways by aerobic methylotrophic bacteria.A. NAD-linked formaldehyde dehydrogenase (FALD), B. Glutathione (GSH)-linked FALD, C. Mycothiol (MySH)-linked FALD, D. Folate-linked pathways, (upper) involving THMPT and (lower) involving THF(Lidstrom, 2006).

Formate Dehydrogenase

Formate is the next oxidation step of methanol after formaldehyde. As can be seen from Figure 1.1, no methylotrophic organism assimilates methanol carbon at the level of formate, but formate represents a valuable source for recycling cofactors ($NAD(P)^+$ to NAD(P)H) and also for energy generation. Under anoxic conditions formate, produced from pyruvate by the enzyme pyruvate formate-lyase (PFL), serves as a major electron donor for a variety of inducible respiratory pathways that use terminal acceptors other than molecular oxygen in *E. coli* (Unden and Bongaerts, 1997). Formate is also an important branch point in the anaerobic Wood-Ljungdahl pathway. Carbon dioxide is reduced to formate by FDH before the enzymes mentioned above (see Formaldehyde oxidation section) can react with formate.

Formate dehydrogenase (FDH) is one of the most extensively studied microbial enzymes. It catalyzes the two electron oxidation of formate to carbon dioxide. It also reduces $NAD(P)^+$ to NAD(P)H during this reduction. Because of this reduction, it is one of the most frequently used enzymes in dehydrogenase-based synthesis of optically active compounds (Tishkov and Popov, 2006). The availability of formate as a cheap substrate also makes FDH an enzyme of choice. There have been extensive efforts to engineer FDHs from different microorganisms to have a higher catalytic activity and increased stability (Tishkov and Popov, 2006). Most microorganisms have been shown to carry multiple copies of formate dehydrogenases. *E. coli* has three distinct FDH isoenzymes (Sawers, 1994). *M. extorquens*AM1 has four functional FDHs (Chistoserdova et. al., 2004 and 2007) although only one has been shown to be indispensable for growth on methanol (2007); mutations in *fdh4A* resulted in a phenotype that cannot grow on methanol. *Methylobacillusflagellatus* is another microorganism that possesses more than one formate dehydrogenase. Like *M. extorquens* AM1, only one of the genes was crucial for growth of *M. flagellatus* (Hendrickson et. al., 2010).

Also, if nitrate is present in the medium when *E. coli* is grown anaerobically, a formate-nitrate respiratory chain may be formed by FDHs with nitrate reductase as a terminal electron acceptor (Enoch and Lester, 1974, 1975). Fdh-N of *E. coli* is part of nitrate reduction where under anaerobic conditions formate oxidation is coupled to nitrate reduction via a lipid-soluble quinone (Jormakka*et. al*., 2003). The capability of FDHs functioning in the formate-nitrate respiratory chain may bestow methylotrophic bacteria with the ability to grow in anaerobic environments.

Molybdenum and tungsten are the most abundant metals found in FDHs, although selenium may also present in some FDHs. The FDH from the sulfate reducing bacterium*Desulfovibrio alaskensis* NCIMB 13491 was shown to contain either molybdenum or tungsten (Brondino et. al., 2004). In *Methylobacterium*sp. RXM, the addition of molybdenum (0.6 µM) and tungsten (0.6 to 0.9 µM) to shake flask cultures yielded the highest formate dehydrogenase activity but these two metals were found to have antagonistic effects on the formate dehydrogenase enzyme that contains the other metal (Girio et. al., 1998). The FDH-H of *E. coli*was shown to have selenium in the active site of the enzyme, and to carry out oxidation of formate into CO_2 without an oxygen transfer (Khangulov et al., 1998).

Anaerobic methylotrophic microorganisms

The long held belief that some microorganisms such as *Bacillus subtilis* are strictly aerobic has been changed recently (Nakano and Zuber, 1998). *B. subtilis* has been shown to grow under an anaerobic atmosphere where nitrate or nitrite is present as an electron acceptor. Similarly, another bacillus, *B. macerans* was shown to respire anaerobically in the presence of nitrate or nitrite with glycerol as an electron donor (Glaser et. al., 1995). Although the presence and functionality of nitrogenase genes in methane oxidizing microorganisms may have suggested the possibility of anaerobic growth (Oakley and Murrel, 1988), methylotrophic microorganisms were not demonstrated to grow anaerobically until recently (Raghoebarsing et. al., 2006; Ettwig et. al., 2008).

Researchers have isolated methylotrophic microorganisms from environments where oxygen is thought to be scarce or absent. Strains of *Methylobacterium* were isolated from legume nodules and shown to fix nitrogen where the strict absence of oxygen is required (Sy et. al., 2001). Oxygen is thought to be a strict requirement for the oxidation of methanol or methyl amines as the terminal electron acceptor. A terminal electron sink is required in both oxidation cases, either for the removal of electrons from methanol through the electron transport chain in the membrane or by the oxidation of NAD^+ to NADH. Methylotrophs are thought to be strict aerobes because of this requirement. The necessity for oxygen is more profound for oxidation of methane to methanol where molecular oxygen is added to methane. However, recent studies demonstrated that sulfate and nitrate can serve as a source of molecular oxygen under anaerobic conditions (Knittel and Boetius, 2009). This proof of concept study has shown that anaerobic oxidation of methane can be coupled to denitrification (Raghoebarsing et. al., 2006). Although this oxidation was achieved by a consortium of microorganisms that include both Archaea and bacteria, researchers were able to separate the bacteria from the Archaea and demonstrate the anaerobic oxidation of methane by bacteria (Ettwig et. al., 2008).

Genetic manipulations of methylotrophs

The development of genetic and molecular biology techniques available for use in methylotrophs traditionally lagged behind those of other microorganisms. The nature of obligate methylotrophy and the limited number of antibiotic markers that can be employed in methylotrophs accounted for this lag. The published reports of the genome sequences of model methylotrophic organisms like *M. capsulatus* (Bath) and *M. extorquens* AM1 lagged behind reported sequences of other microorganisms and only recently became available (Chistoserdova et. al., 2009 and Ward et. al., 2004). Triparental or biparental conjugative mating techniques have been preferred as a method of genetic transformation rather than techniques like electroporation. The *E.coli* strain S17-1 was successfully used as donor in mating experiments with Gram-negative bacteria (Simon et. al., 1983). Matings are usually carried out on unselective solid media (Ali et. al., 2006, Smith et. al., 2002 and Warner et. al., 1980). Stolyar and coworkers (Stolyar et. al., 1995) proved the possibility of transfer and expression of the Inc-P incompatibility group in cells of obligate methane-oxidizing bacteria *Methylomonasrubra* 15sh and *Methylococcus thermophilus* 111p. They also showed the intraspecific gene transfer in methane-oxidizing bacteria for the first time. The cloning and usage of the methanol-regulated promoter sequence from the methanol oxidation gene *moxF* found in facultative methylotrophic bacterium *Methylobacteriumorganophilum* XX has made the homologous and heterologous expression in methylotrophic bacteria possible (Choi et. al., 2006 and Xu et. al., 1993). However, expression of methane monooxygenase from *M. capsulatus* (Bath) in *E.coli* did not yield a fully functional enzyme (West et. al., 1992). Promoter-probe vectors with reporter genes encoding green fluorescent protein (GFP), catechol 2,3-dioxygenase (XylE) or β-galactosidase (LacZ) were also designed and used in the methylotroph *M. capsulatus* (Bath) (Ali and Murrel, 2009). Methane monooxygenase mutants of *Methylosinustrichosporium* have been successfully constructed by marker-exchange mutagenesis (Martin and Murrel, 1995).

Bacillus methanolicus

B. methanolicus is a Gram-positive aerobic methylotroph growing at temperatures between 35 and 60 °C. Various wild type strains that are able to grow on methanol as a sole carbon source were isolated by different researchers (Arfman et. al., 1992). Wild-type strains of *B. methanolicus*isolated at the University of Minnesota have been reported to secrete 58 g/l of L-glutamate in fed-batch cultures (Schendel et. al., 2000) while classical mutants can secrete 37 g/l of L-lysine (Hanson et. al., 1996), at 50 °C. The ribulose monophosphate (RuMP) pathway is used to assimilate the carbon derived from methanol. Methanol is converted to formaldehyde by a methanol dehydrogenase (MDH) which is encoded by a gene carried on a 19 kb endogenous plasmid (Brautaset et. al., 2004). The MDH in *B. methanolicus* is shown to be an NAD-dependent, decameric protein with 43,000-Mr subunits. Each subunit contains a tightly bound NAD(H) molecule, one zinc and one or two Mg ions

(Arfman et. al., 1997). The only gene described encoding MDH was reported to be on the pBM19 plasmid (Brautaset et al., 2004).

This plasmid pBM19 also has additional genes, encoding enzymes of the RuMP pathway, *glpX*, *fba*, *tkt*, *pfk* and *rpe*, which encode the enzymes fructose-1,6-biphosphatase (FBPase), fructose-1,6-biphospate aldolase (FBPA), transketolase, phosphofructokinase, and Ru-5-P-3 epimerase, respectively. Five of the corresponding genes of RuMP pathway enzymes are carried on the pBM19 along with *mdh*. However, the genes *hps* and *phi*, encoding the first two enzymes of the RuMP pathway, hexulose phosphate synthase and phosphohexuloisomerase, are located on the chromosome. The overexpression of these genes resulted in reported increased tolerance to higher formaldehyde concentrations in shake flask studies compared to wild type strains (Jakobsen et. al., 2006). There had been attempts to increase the amino acid production by overexpression of the RuMP pathway or TCA cycle enzymes. The overexpression of aspartokinase increased the L-lysine production while overexpression of pyruvate carboxylase did not have a positive effect on the rate of L-lysine or L-glutamate production (Brautaset et. al., 2010 and Jakobsen et. al., 2009).

Genetic manipulations in *B. methanolicus* were made possible by the pioneering work of Cue *et. al.,*1997: *B. methanolicus* was successfully transformed by protoplast formation for the first time. *E.coli-B.methanolicus* shuttle vectors, pDQ507 and pDQ508 were successfully constructed and shown to be stable in *B. methanolicus* up to 60 generations even in nonselective conditions. The stability of *E.coli-B.subtilis* shuttle vector pHP13 was also assessed. All three plasmids were shown to replicate by a rolling-circle mechanism which contributed to their stability in *B. methanolicus*. Methylation of the restriction site *Bme*TI, isoschizomer of *Bcl*I, also added stability to the transformed plasmids, since *B. methanolicus* was shown to have the *Bme*TI restriction-modification system (Cue et. al., 1996).

The dissimilation of carbon from formaldehyde via conversion into formate and then carbon dioxide, which leads to loss of carbon, has been shown for *B. methanolicus* MGA3 by ^{13}C NMR and isotope-ratio mass spectrometry (Pluschkell and Flickinger, 2002). However, the difference in carbon dissimilation rates of strains MGA and PB1, as shown in Figure 1.4, suggests that there should be significant difference in how these two strains dissimilate the carbon (Gutmacher, 2009).

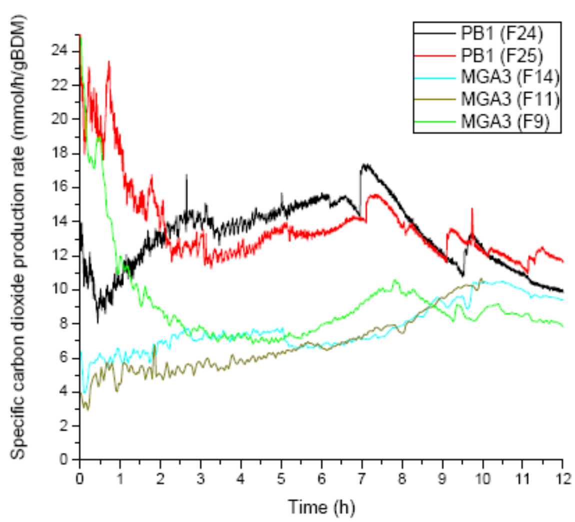


Figure 1.4. Specific CO_2 production rates of *B. methanolicus* MGA3 and PB1 in a 1-L fed-batch fermentation (Gutmacher, 2009).

The presence of formaldehyde and formate dehydrogenase enzymes was shown by enzyme assays of the crude cell extracts in Minnesota (Smith, 2007), but were never shown on the genetic level. Candidate genes were cloned from *B. methanolicus* PB1 by a research group in Bath, U.K. using degenerate primers but the expression of these clones did not confer any formaldehyde or formate dehydrogenase activity either in*E. coli*or in *B. subtilis*(Mozzanega, 2010). Furthermore, the expression of plasmid-borne *mdh* in *E. coli* or *B. subtilis*carried out by this group did not yield an active enzyme in the same study (Mozzanega, 2010). The lack of activity from the *mdh* enzyme raises the question about the validity of the enzyme assays Mozzanega conducted since the enzyme was characterized before and the activity was shown clearly (Arfman et. al., 1997). Identification of a thermo-stable formaldehyde dehydrogenase and formate dehydrogenase from *B. methanolicus* may represent a significantly improved enzyme for use as biocatalysts by the chemical industry (Schrader et. al. 2009).

Recently, the genome sequences of both *B. methanolicus* MGA3 and PB1 were published ahead of us by our former collaborators in Norway (Heggeset *et. al.,* 2012). The difference of *B. methanolicus* strains MGA3 and PB1 in terms of carbon dissimilation rate is confirmed in the same report. This report summarizes the methanol assimilation and dissimilation pathways of *B. methanolicus* as depicted in Figure 1.5 along with key findings including possible growth on glucose. Some of these findings were already written in Chapter 3 before the report was published. The Norwegian report is presented as Appendix C.

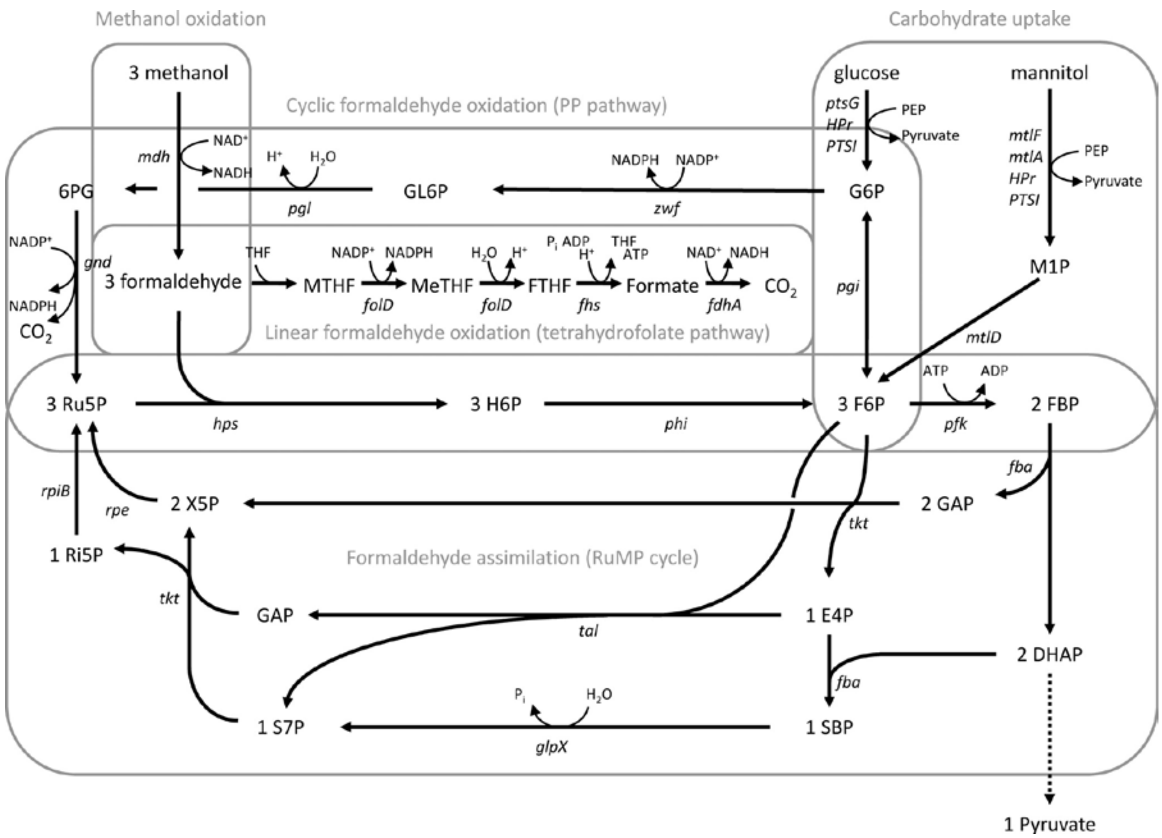


Figure 1.5. Proposed methylotrophic and important carbohydrate assimilation pathways in *B. methanolicus* (Heggeset *et. al.*, 2012)

Thesis Goals

The goals of this thesis are to:

- Investigate the tolerance of *B. methanolicus* for methanol, formaldehyde, and formate
- Sequence the *B. methanolicus* MGA3 genome, annotate the draft sequence.
- Depict the methanol dissimilation pathway from the annotated draft genome sequence

- Elucidate the methanol and formaldehyde assimilation and dissimilation pathways of *B. methanolicus* strains MGA3 and PB1
- Investigate the regulation of formaldehyde dissimilation pathway genes
- Characterize the enzymes involved in formaldehyde assimilation and dissimilation
- Compare the carbon dioxide evolution of different *B. methanolicus* strains in a small methanol-fed bioreactor
- Clone, characterize and compare methanol, formaldehyde and formate dehydrogenases of different bacilli with those of *B. methanolicus*
- Investigate the possible anaerobic growth of *B. methanolicus* on different media and gases

CHAPTER 2

GROWTH CHARACTERISTICS OF *B. methanolicus*

Introduction

Various wild type strains of *B. methanolicus* have been described by different research groups from various environments (Dijhuizen et. al., 1988, Arfman et. al., 1992, Schendel et. al., 1990). Most of these strains are restricted facultative methylotrophs in that they can also grow on a limited number of sugars like maltose or mannitol (Arfman et. al., 1992). The growth energetics of methylotrophs differs according to the pathway they employ to fix the formaldehyde produced from methanol. Extensive theoretical modeling studies have been conducted to estimate the growth yield from methanol and its metabolites (Anthony, 1978 and Van Dien and Lidstrom, 2002). The work of Anthony (1978) modeled the growth yield of different variants of RuMP and serine pathway methylotrophs based on the Y_{ATP} while Dijkhuizen et. al. (1992) summarized (see Introduction) the energetics of formation of pyruvate from formaldehyde by different variants of the RuMP pathway. In addition, Van Dien and Lidstrom modeled the growth of *M. extorquens* AM1 (Serine pathway) and *Methylobacillus falgellatus* KT (RuMP pathway). Anthony (1978) found that 1.5 moles of NADH, 29 moles of ATP and 3 moles of NH_3 are required to synthesize 306 g of biomass from 12 moles of formaldehyde molecules by the FBPA variant of the RuMP pathway, while 17.5 moles of NADH, 41 moles of ATP, 4 moles of CO_2, 3 moles of NH_3, and 4 moles of O_2 are needed to synthesize the same amount of biomass from 8 moles of formaldehyde via the *icl*$^-$ variant of serine cycle. These models have important implications for the biotechnological use of methylotrophs.

The biotechnological aspects of methylotrophs have been reviewed by Schrader et. al. (2009), Trotsenko et. al. (2005), and Brautaset et. al. (2007). Methylotrophs are promising candidates for the production of forage proteins for livestock, production of Poly-β-hydroxybutyrate/valerate, exopolysaccharides, ectoine, phytohormones and vitamins (Trotsenko et. al., 2005). They also contain high levels of various dehydrogenase enzymes that may be used in generation of reduced compounds like NADH. The RuMP pathway methylotrophs are superior to serine pathway microorganisms in that they produce almost two fold higher cell yield per gram of methanol carbon (Schrader et al., 2009). Another advantage of the RuMP pathway organisms is the higher biomass productivity; RuMP pathway methylotrophs have an average productivity of 28.4 g_{CDM}/l/h, while productivity of serine cycle microorganisms ranges between 1.2-3.6 g_{CDM}/l/h. A portion of methanol is dissimilated as carbon dioxide and lost. In RuMP pathway methylotrophs, approximately 38% of methanol carbon is dissimilated (Schrader et. al., 2009). In this regard, understanding of the formaldehyde dissimilation pathway is equally important as the assimilation pathway since it determines the portion of substrate carbon that is lost to carbon dioxide.

An important limitation of the methylotrophs is the toxicity of the methanol or its metabolites. While this toxicity can lower the contamination risk during the fermentation process, it can also suppress the growth at higher than the optimum concentrations. Thus, any fermentation process is required to control the methanol level. This chapter aims to characterize the toxicity of methanol and its metabolites on the growth of *B. methanolicus* strains MBA3 and PB1. Toxicity of hydrogen peroxide and acetaldehyde were also assessed as oxidative stress inducing compounds.

Materials and Methods

Growth Media: A modified version of the minimal media (MTYM) developed by Dijkhuizen et al. (1988) was used as growth media. It contains 10.9 mM $NaH_2PO_4 \cdot H_2O$ (1.5g/l), 23.5 mM K_2HPO_4 (4.1 g/l), 15.9 mM $(NH_4)_2SO_4$(2.1 g/l), and 0.25g/l of yeast extract, pH 7.0. The trace metal stock solution (Table 2.1) and vitamin stock (Table 2.2.) solution were added to the final concentration of 1X. The trace metal solution was sterilized by autoclaving for 20 min and vitamin stock solution was filter sterilized with a 0.22µ filter. The $MTYM_{MeOH}$ medium contained 150 mM of methanol added to the medium while $MTYM_{MAN}$ medium contained 55 mM mannitol (10.02 gr/L), unless otherwise stated. Chloramphenicol is added to the media when required at the concentration of 5µg/l.

Table 2.1. Composition of trace metal solution (1000X)

Ingredient	**F.W. g/mol**	**Stock 1000x mM**	**Stock 1000x g/L**
Citrate·H_2O	210.14	124.4	26.14
$MgCl_2 \cdot 6\ H_2O$	203.29	1000	203.3
$MnCl_2 \cdot 4\ H_2O$	197.91	100	19.8
$CaCl_2 \cdot 2\ H_2O$	147.02	100	14.7
$FeSO_4 \cdot 7\ H_2O$	278.02	40	11.12
$ZnCl_2$	126.29	2	0.27
H_3BO_3	61.84	1	0.06
$Na_2MoO_4 \cdot 2\ H_2O$	241.98	0.4	0.1
$CoCl_2 \cdot 6\ H_2O$	237.9	0.34	0.08
$CuCl_2 \cdot 2\ H_2O$	170.49	0.32	0.05

Bacterial stratins: *B. methanolicus* MGA described by Schendel et al. (1990) was germinated from spores as follows; 50 µl of spore stock was incubated at 80 °C for 15 min and then transferred into a pre-warmed $MTYM_{MeOH}$ media and incubated at 50 °C, 270 rpm in a shaker. *B. methanolicus* PB1 (ATCC 51375), first isolated and described by Arfman et. al. (1992), was purchased from the ATCC. The cryoculture stock was prepared from the exponentially growing *B. methanolicus* culture with 10% final glycerol concentration and by flash freeze in liquid nitrogen.

Table 2.2. Composition of vitamin stock solution (1000X)

Vitamin	**F.W. g/mol**	**Stock 1000x mM**	**Stock 1000x mg/L**
d-biotin (B_7)	244.31	0.409	100
thiamine·HCl (B_1)	337.27	0.296	100
riboflavin (B_2)	376.37	0.266	100
pyridoxine·HCl (B_6)	169.18	0.591	100
pantothenic acid (B_5)	219.24	0.456	100
nicotinic acid (B_3)	123.11	0.812	100
p-aminobenzoic acid (B_x)	137.14	0.146	20
folic acid (B_9)	441.41	0.023	10
cyanocobalamin (B_{12})	1355.39	0.007	10
lipoic acid	206.33	0.048	10

Methanol and *B. methanolicus* growth:*B. methanolicus* was grown in 50 ml $MTYM_{MeOH}$ medium in 250 ml baffled shaker flasks at 50 °C and 270 rpm in a shaker (HT-INFORS) for 10-11 h and inoculated into pre-warmed MTYM media with methanol concentrations ranging between 6 to 2000 mM. The specific growth (μ) rate is calculated from the exponential phase of the growth. The growth of *B. methanolicus* in $MTYM_{MeOH}$ with 25 mM nitrate, 2X YE and without YE were also evaluated.

***B. methanolicus* growth on mannitol:** Two hundred fifty microliters of *B. methanolicus* was inoculated into $MTYM_{MAN}$ medium supplied with 1X vitamin and trace metal stock solution and grown for 10-11 hours. Then enough volume was transferred into 50 ml pre-warmed $MTYM_{MAN}$ medium in 250 ml baffled shaker flask to have an OD_{600} ~0.05, with mannitol concentrations of 5, 10, 25 and 55 mM. The specific growth rate (μ) was calculated from exponential phase of the growth.

***B. methanolicus* growth on glucose:** A tube of cryostock was thawed and inoculated into MTYM as media as described above and mannitol was replaced with 50 mM of glucose. The culture was incubated at 50 °C overnight and enough volume was transferred into 50 ml pre-warmed MTYM medium in 250 ml baffled shakers flask to have an OD_{600} ~0.05, with glucose concentrations of 5, 10, and 20 mM. The specific growth rate (μ) was calculated from exponential phase of the growth.

Growth response to different methanol metabolites: Three hundred microliter of frozen *B. methanolicus* is inoculated into 50 ml of $MTYM_{MeOH}$ or $MTYM_{MAN}$ medium in 250 ml baffled shaker flasks and grown 10-11 hours. Then, enough volume was transferred into a 50 ml pre-warmed $MTYM_{MeOH}$ or $MTYM_{MAN}$ to have an OD_{600} ~0.05. The culture was sampled every hour for OD_{600} measurements and spiked with formaldehyde and formate at the concentrations of 0.5, 1, and 2 mM when the OD_{600} of culture is 0.250-0.350. Formaldehyde was prepared fresh from paraformaldehyde as 2M stock in a serum bottle and autoclaved. Formate was prepared from sodium formate in distilled water as 2M stock. For the methanol spike in mannitol-grown cells, the culture was spiked with 100 mM methanol when the OD_{600} was ~0.25.

Growth response to hydrogen peroxide and acetaldehyde: As described above, *B. methanolicus* was grown either on $MTYM_{MeOH}$ or $MTYM_{MAN}$ and spiked with 0.5, 1 and 2 mM hydrogen peroxide and acetaldehyde. The growth was monitored by measuring OD_{600} every hour.

Detection of formaldehyde in the media: A colorimetric method originally developed by Chrastil and Wilson (1975) and modified by Topp and Knowles (1984) was used to measure the formaldehyde concentration in the culture. Briefly; 1 ml of sample was added into 1 ml of tryptophan-ethanol reagent and then, 1 ml of concentrated (90%) H_2SO_4 was added immediately. The mixture was agitated with a vortex mixer and 200 µl of 0.2% $FeCl_3$ (wt/v) was added for color development. Then, mixture was incubated for 1 hour at 70 °C in an oven. The color development was measured at 575 nm after the samples cooled. A standard curve was obtained by diluting formaldehyde stock solution (2M) in MTYM to have a concentration range of 10 to 700 µM in MTYM. No attempt was made to detect formate in the culture.

Detoxification of formaldehyde by methanol and mannitol grown *B. methanolicus* MGA3: *B. methanolicus* MGA3 was grown to OD_{600} of 0.8 in 50 ml of either $MTYM_{MeOH}$ or $MTYM_{MAN}$ in a 250 ml shaker flask and then formaldehyde was added to final concentration of 2 mM. The formaldehyde concentration in the culture media was measured via the method described above by using 1 ml of culture at the time points of 5, 10, 20 and 60 minutes after formaldehyde addition.

Results

Methanol and *B. methanolicus* growth: The growth rate of *B. methanolicus* MGA3 grown on different methanol concentrations is shown in Figure 2.1. The maximum specific growth rate (μ) was achieved (0.65±0.01 h^{-1}) when *B. methanolicus* MGA3 is grown in the MTYM with 10 mM methanol. The growth rate decreased when the methanol concentration is higher than 25 mM, from 0.65±0.007 h^{-1}at 25 mM to 0.23±0.004 h^{-1} at 2 M. It is also noteworthy that *B. methanolicus* MGA3 culture growing in 2 M methanol achieved a final OD_{600} of 3.76±0.09. Figure 2.2 displays *B. methanolicus* MGA3 culture grown in $MTYM_{MeOH}$ and spiked with methanol concentrations of 10, 50 and 100 mM. As seen from the figure, there is no change in the growth of *B. methanolicus* MGA3 growing on the methanol when challenged with the mentioned methanol concentrations. However, a significant growth disturbance is seen with the same methanol concentrations when $MTYM_{MAN}$-grown *B. methanolicus* MGA3 is challenged (figure 2.3).

Mannitol and *B. methanolicus* growth: The growth rate of *B. methanolicus* MGA3 in different mannitol concentrations is shown in Figure 2.4. The specific growth rate (μ) is relatively constant for different mannitol concentrations, 0.53±0.002. As shown in the figure 2.4, there is no decrease or increase in the growth rate depending on the mannitol concentration.

Glucose and *B. methanolicus* growth: Although it was suggested by Heggeset et. al. (2012), growth of *B. methanolicus*on glucose is demonstrated for the first time in this study.

The growth rate of *B. methanolicus* MGA3 in different glucose concentrations is shown in Figure 2.5. The specific growth rate does not change with the glucose concentrations tested, 0.336±0.003 h^{-1} which is considerably lower than growth rate of *B. methanolicus* on mannitol.

Growth response of *B. methanolicus* to formaldehyde and formate: The exponentially growing culture of *B. methanolicus* was spiked with different concentrations of formaldehyde and formate to elucidate the effect of these methanol metabolites on the growth of the microorganism. Figure 2.6 demonstrates the *B. methanolicus* MGA3 grown on methanol and spiked with different concentrations of formaldehyde. As seen from the figure, *B. methanolicus*MGA3 is able to cope with 0.5 mM formaldehyde, but 1 mM formaldehyde almost completely halts the growth of the microorganism, and 2 mM formaldehyde is enough to kill the culture. However, mannitol grown *B. methnaoliucus* MGA3 is more tolerant to formaldehyde. As shown in Figure 2.7, *B. methanolicus* MGA3 is able tolerate even 2 mM formaldehyde.

Figures 2.8 and 2.9 demonstrate the response of *B. methanolicus* MGA3 grown in $MTYM_{MeOH}$ and $MTYM_{MAN}$ respectively, and spiked with the same concentrations of formate as was formaldehyde. As seen from the figures, formate at the spiked concentrations does not interfere with*B. methanolicus* MGA3 growth.

Growth response of *B. methanolicus* to acetaldehyde and hydrogen peroxide: The growth response of *B. methanolicus* MGA3 to acetaldehyde, to compare with the stress response of formaldehyde and hydrogen peroxide as general oxidative stress agents, was tested. As can be seen from the figures 2.12 and 2.13, acetaldehyde with the tested concentrations does not interfere with growth of *B. methanolicus* MGA3, grown in either $MTYM_{MeOH}$ or $MTYM_{MAN}$. However, figures 2.10 and 2.11 demonstrate that even 0.5 mM hydrogen peroxide was enough to halt the growth of *B. methanolicus* MGA3 completely, whether the cells are grown on methanol or mannitol.

Detoxification of formaldehyde by methanol and mannitol grown *B. methanolicus* MGA3: The standard curve obtained with known formaldehyde concentrations in MTYM medium is shown in figure 2.14. A linear curve with the R-square value of 0.997 was obtained with the formaldehyde range of 10 to 700 µM.

Detoxification (removal) of formaldehyde from the media is shown in figure 2.15. *B. methanolicus* MGA3 detoxifies formaldehyde faster when grown in $MTYM_{MAN}$. The concentration of formaldehyde falls below detection limit (10 µM in sample) in 60 minutes when *B. methanolicus* MGA3 is grown in $MTYM_{MAN}$ whereas in $MTYM_{MeOH}$ culture, it is still at the level of 457 µM. The concentrations of formaldehyde in both $MTYM_{MeOH}$ and $MTYM_{MAN}$ grown cultures were below the detection limit prior to the addition.

Conclusion

Toxicity of methanol and its metabolite formaldehyde is one of the major concerns for the biotechnological uses of methylotrophs. *B. methanolicus* gains the highest specific growth rate with methanol concentrations of 10 and 25 mM. The specific growth rate decreases in methanol concentrations higher than 25 mM. The response of $MTYM_{MAN}$ grown cells to methanol spikes suggests that methanol or its metabolites are toxic to *B. methanolicus* MGA3 even at 50 mM which is lower than $MTYM_{MeOH}$. Considering this with the specific growth rate in $MTYM_{MeOH}$ being lower than maximum, it would not be misleading to state that growth of *B. methanolicus* MGA3 in $MTYM_{MeOH}$ is stressful and *B. methanolicus* MGA3 successfully adapted to this stress albeit with lower specific growth rate. The addition of 100 mM methanol into the culture growing in $MTYM_{MeOH}$ did not produce repression of the growth due most probably to adaptation to stress environment. Taking the methanol evaporation rate (Nilasari et. al., 2012) and growth of *B. methanolicus* MGA3 into account, the maximum methanol concentration in $MTYM_{MeOH}$ does not exceed 250 mM. The growth rates at 150 mM and 300 mM methanol concentrations are comparable, 0.55±0.007 and 0.52±0.01, respectively.

Formaldehyde is an important branch point of methanol metabolism in that carbon of formaldehyde is either assimilated into biomass or lost into the air as carbon dioxide. Because of its toxicity, the intercellular concentration of formaldehyde needs to be controlled to avoid any toxic effects. As seen from figures 2.6, formaldehyde is toxic to *B. methanolicus* even at the 1 mM concentration in the culture supernatant when *B. methanolicus* MGA3 is grown in $MTYM_{MeOH}$.However, when grown in $MTYM_{MAN}$ (figure 2.7), even 2 mM formaldehyde can be tolerated by *B. methanolicus* MGA3. This is most probably due to saturation of formaldehyde detoxification pathways in methanol grown cells, since mannitol grown cells are also able to remove formaldehyde from the media faster (figure 2.15). One route of formaldehyde removal is the assimilation into biomass via the RuMP pathway. Jakobsenet. al., (2006) demonstrated the upregulation of RuMP pathway genes on both chromosome and plasmid. Another possible route of formaldehyde removal is the dissimilation into carbon dioxide through oxidation to first formate and then carbon dioxide. This pathway is also important since it determines the oxidative capacity of the microorganism. Microorganisms with higher oxidative capacity require more effective cooling due to increased heat output (Schrader et. al., 2009) beside the loss of substrate carbon into air.

Formate is the next compound after formaldehyde in the dissimilation pathway. Addition of formate into $MTYM_{MeOH}$ or $MTYM_{MAN}$ grown cells did not produce any growth inhibition as can be seen from the figures 2.8 and 2.9. This is most probably due to low concentrations of formate being added since Pluschkell (1998) showed the inhibition with 25 mM formate. The accumulation of formaldehyde and formate in continuous *B. methanolicus* MGA3 culture when spiked with methanol was also shown by Pluschkell and Flickinger (2002), and formaldehyde and formate dehydrogenase activity in the crude cell extract was shown by Smith (2007). However, no formaldehyde or formate dehydrogenase genes have been identified and characterized from *B. methanolicus* MGA3. Identification of these genes and elucidation of their regulation are important to understand formaldehyde dissimilation in *B. methanolicus*.

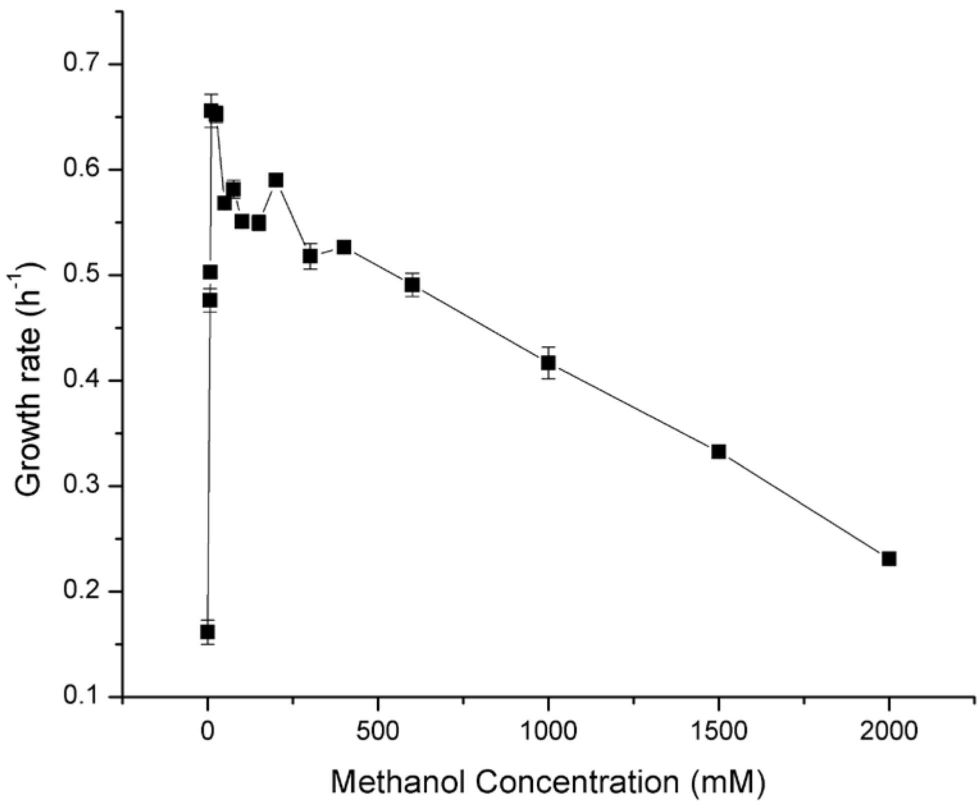


Figure 2.1. The maximum specific growth rate of *B. methanolicus* MGA3 in different methanol concentrations (n=3). *B. methanolicus* MGA3 was grown in 50 ml MTYM medium in 250 ml shaker flasks with different methanol concentrations. The µ was calculated from the exponential phase of the growth.

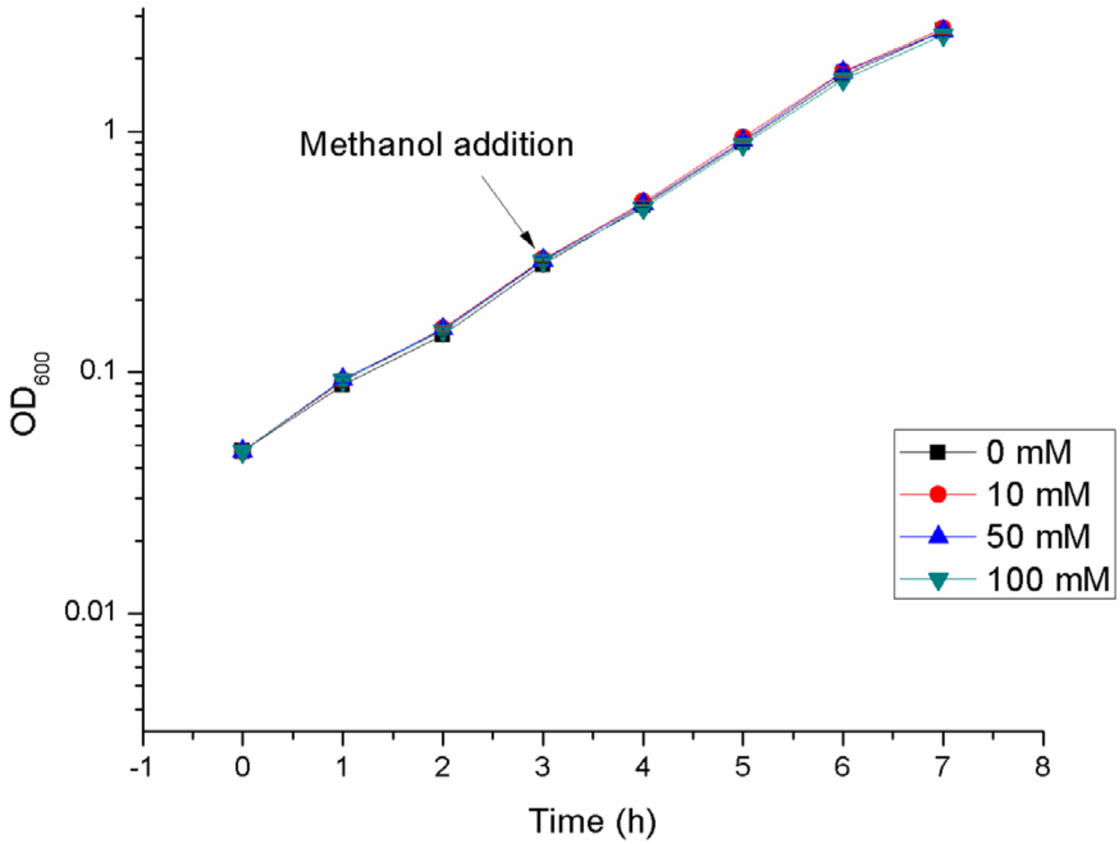


Figure 2.2. Response of $MTYM_{MeOH}$ grown *B. methanolicus* MGA3 to methanol spikes.*B. methanolicus* MGA3 was grown in a 50 ml $MTYM_{MeOH}$ in 250 ml baffled shaker flasks and challenged with different concentrations of methanol (n=3) when the cells were growing exponentially (OD_{600}~0.25) .

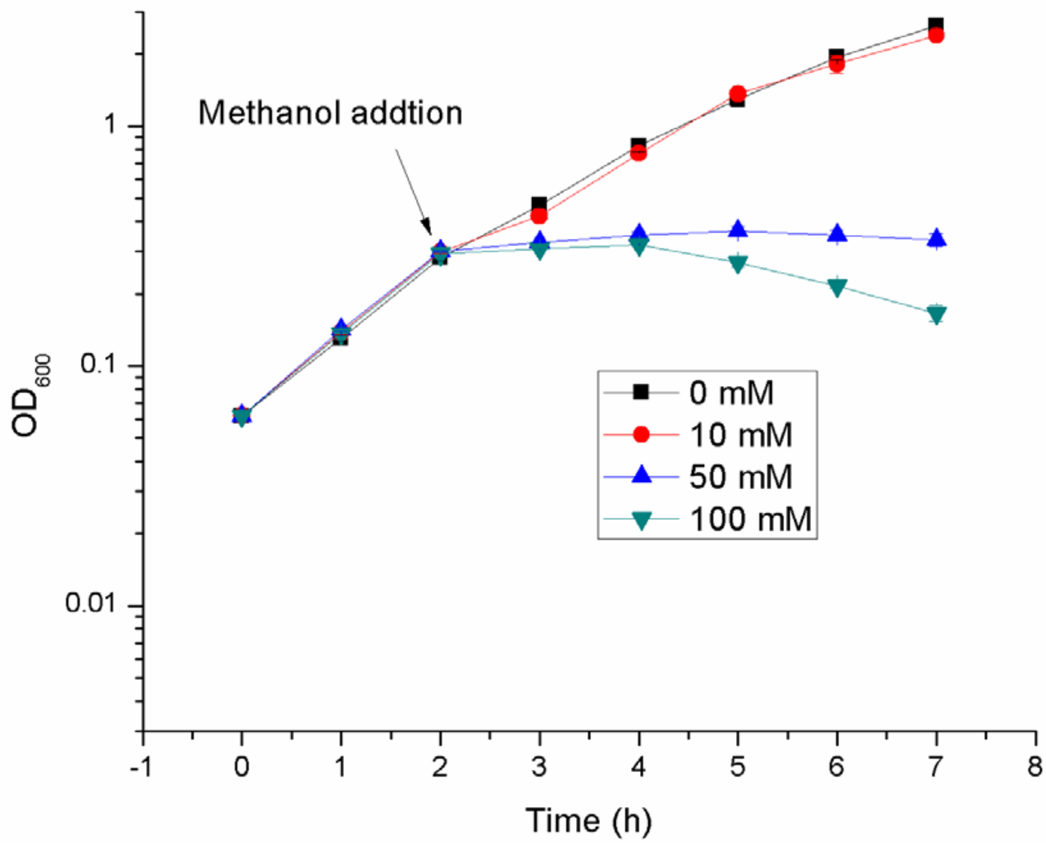


Figure 2.3. Response of $MTYM_{MAN}$ grown *B. methanolicus* MGA3 to methanol spikes.*B. methanolicus* MGA3 was grown in a 50 ml $MTYM_{MAN}$ in 250 ml baffled shaker flasks and challenged with different concentrations of methanol (n=3) when the cells were growing exponentially (OD_{600}~0.25) .

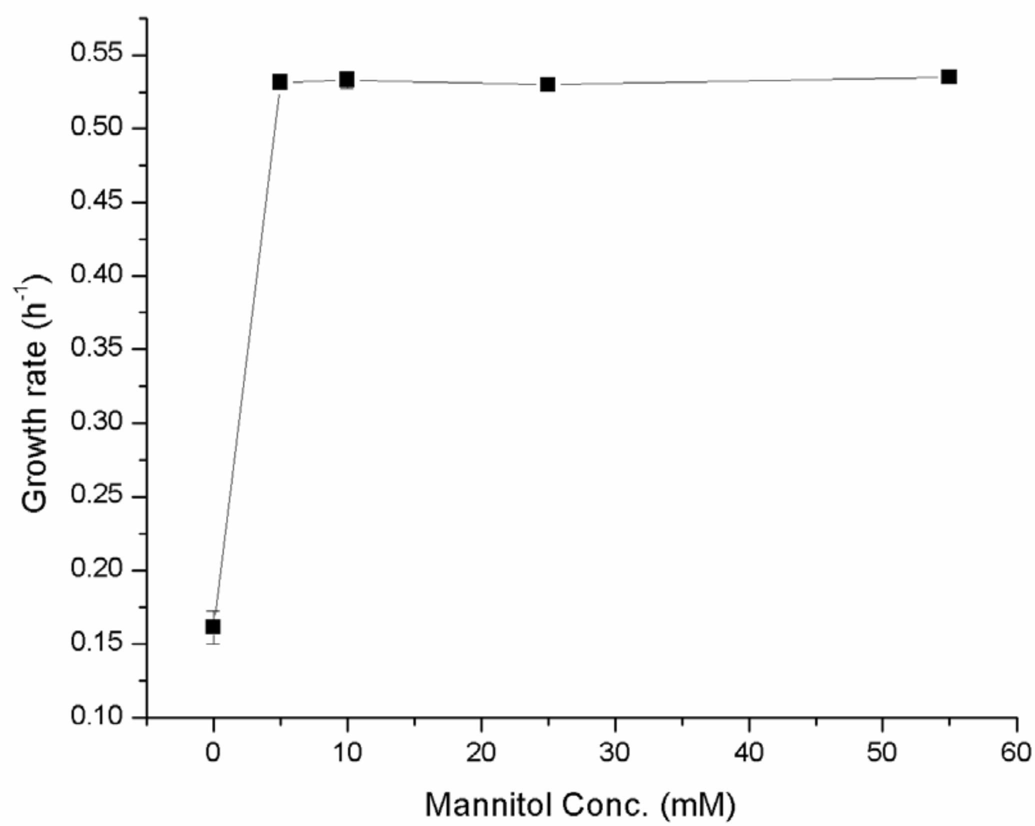


Figure 2.4. The maximum specific growth rates of *B. methanolicus* MGA3 on different mannitol concentrations (n=3). *B. methanolicus* MGA3 was grown in MTYM medium with different mannitol concentrations and μ was calculated from exponential phase of the growth.

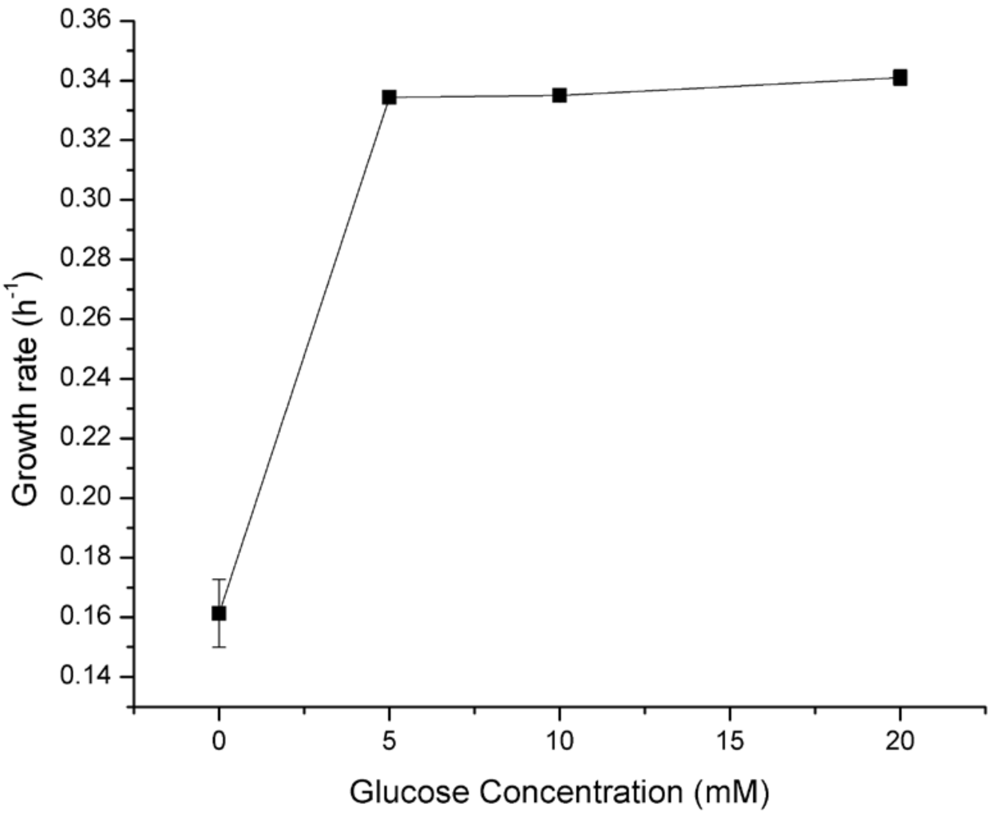


Figure 2.5. The maximum specific growth rates of *B. methanolicus* MGA3 on different glucose concentrations (n=3). *B. methanolicus* MGA3 was grown in MTYM medium with different glucose concentrations and μ was calculated from the exponential phase of the growth.

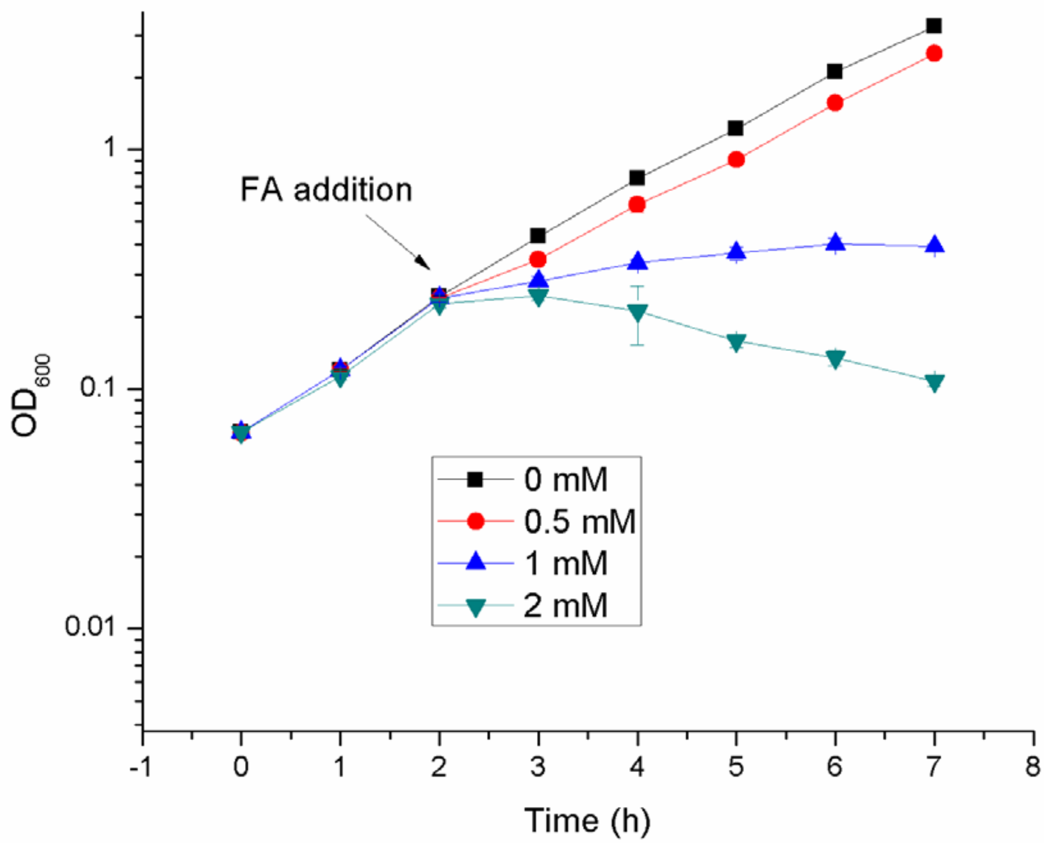


Figure 2.6. Response of $MTYM_{MeOH}$ grown *B. methanolicus* MGA3 to formaldehyde spikes. Formaldehyde was added to the exponentially growing *B. methanolicus* MGA3 culture (OD_{600} ~0.25) and growth was monitored by optical density at 600 nm (n=3).

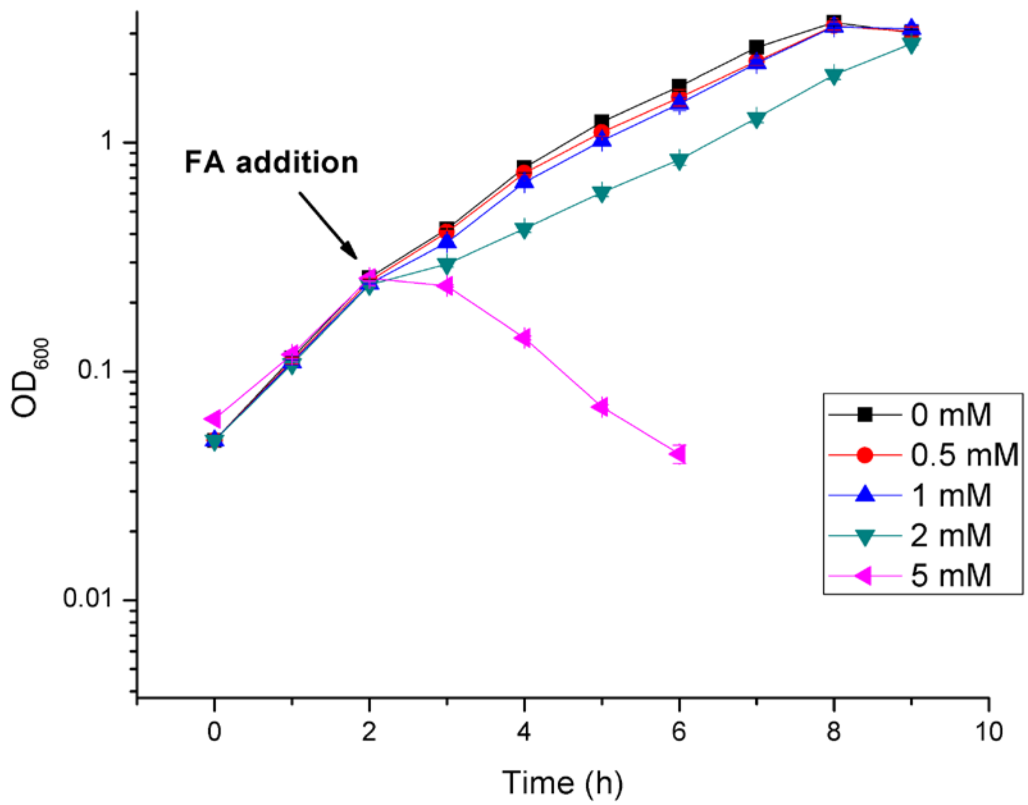


Figure 2.7. Response of $MTYM_{MAN}$ grown *B. methanolicus* MGA3 to formaldehyde spikes. Formaldehyde was added to the exponentially growing *B. methanolicus* MGA3 culture (OD_{600} ~0.25) and growth was monitored by optical density at 600 nm (n=3).

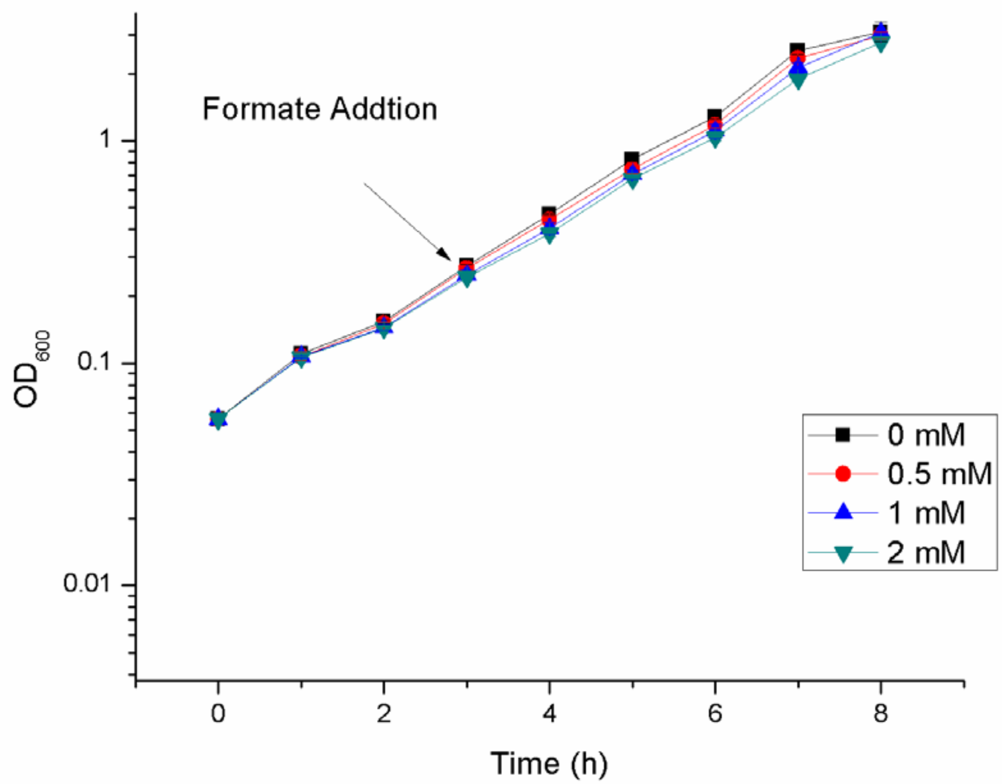


Figure 2.8. Response of $MTYM_{MeOH}$ grown *B. methanolicus* MGA3 to formate spikes. *B. methanolicus* MGA3 was grown in a 50 ml $MTYM_{MeOH}$ in 250 ml baffled shaker flasks and challenged with different concentrations of formate (n=3) when the cells were growing exponentially (OD_{600}~0.25) .

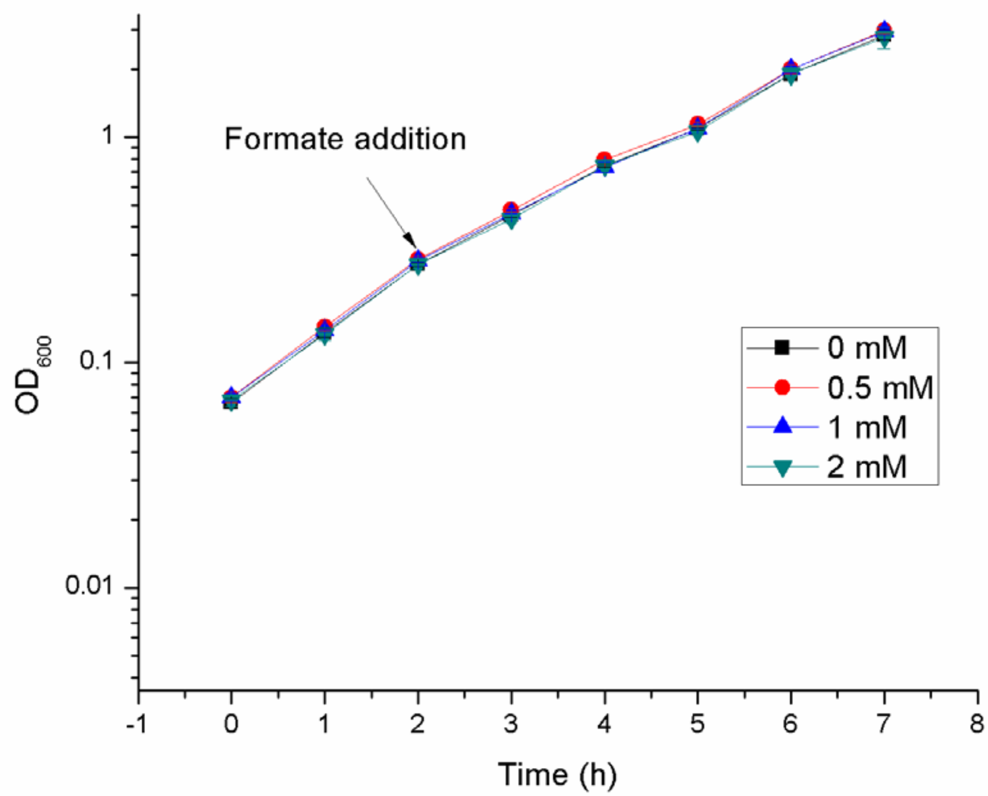


Figure 2.9. Response of $MTYM_{MAN}$ grown *B. methanolicus* MGA3 to formate spikes. *B. methanolicus* MGA3 was grown in a 50 ml $MTYM_{MAN}$ in 250 ml baffled shaker flasks and challenged with different concentrations of formate (n=3) when the cells were growing exponentially (OD_{600}~0.25) .

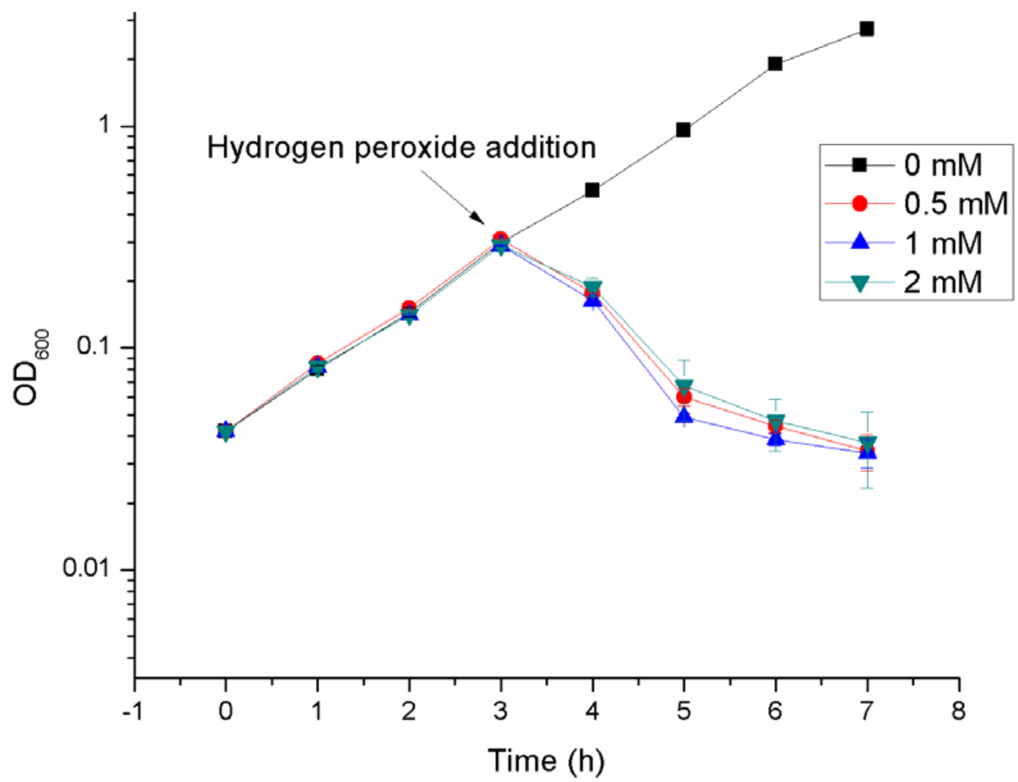


Figure 2.10. Response of $MTYM_{MeOH}$ grown *B. methanolicus* MGA3 to hydrogen peroxide spikes. *B. methanolicus* MGA3 was grown in a 50 ml $MTYM_{MeOH}$ in 250 ml baffled shaker flask and challenged with different concentrations of hydrogen peroxide (n=3) when the cells were growing exponentially (OD_{600}~0.25) .

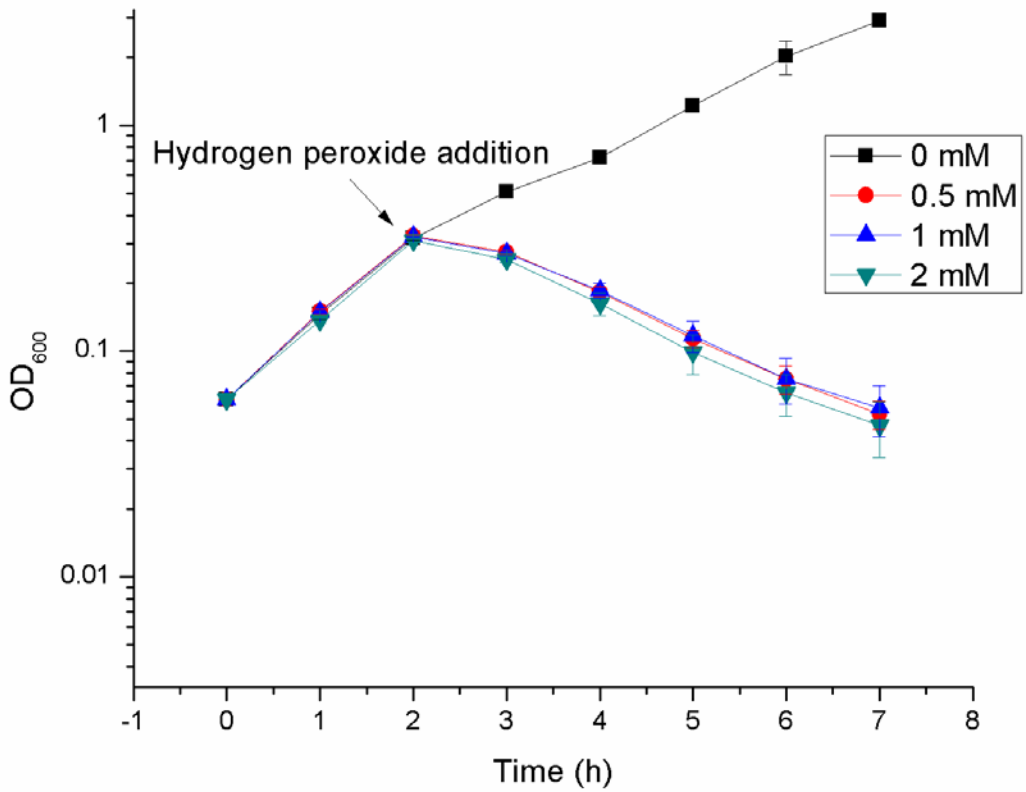


Figure 2.11. Response of $MTYM_{MAN}$ grown *B. methanolicus* MGA3 to hydrogen peroxide spikes. *B. methanolicus* MGA3 was grown in a 50 ml $MTYM_{MAN}$ in 250 ml baffled shaker flask and challenged with different concentrations of hydrogen peroxide (n=3) when the cells were growing exponentially (OD_{600}~0.25) .

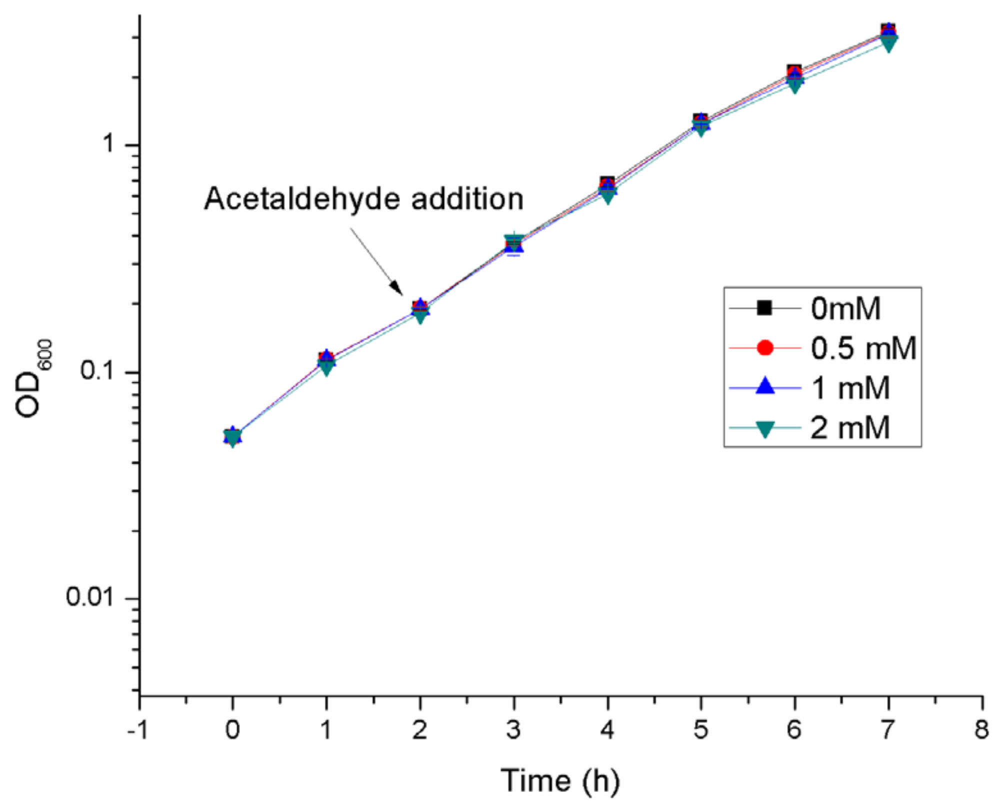


Figure 2.12. Response of $MTYM_{MeOH}$ grown *B. methanolicus* MGA3 to acetaldehyde spikes. *B. methanolicus* MGA3 was grown in a 50 ml $MTYM_{MeOH}$ in 250 ml baffled shaker flasks and challenged with different concentrations of acetaldehyde (n=3) when the cells were growing exponentially (OD_{600}~0.25) .

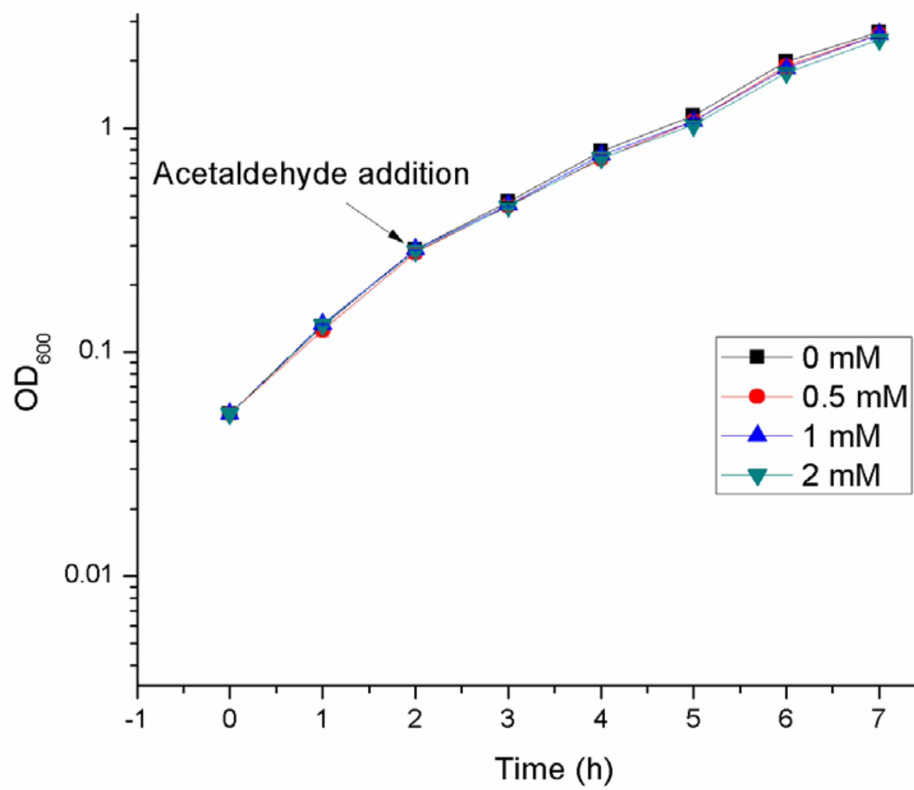


Figure 2.13. Response of $MTYM_{MAN}$ grown *B. methanolicus* MGA3 to acetaldehyde spikes. *B. methanolicus* MGA3 was grown in a 50 ml $MTYM_{MAN}$ in 250 ml baffled shaker flasks and challenged with different concentrations of acetaldehyde (n=3) when the cells were growing exponentially (OD_{600}~0.25) .

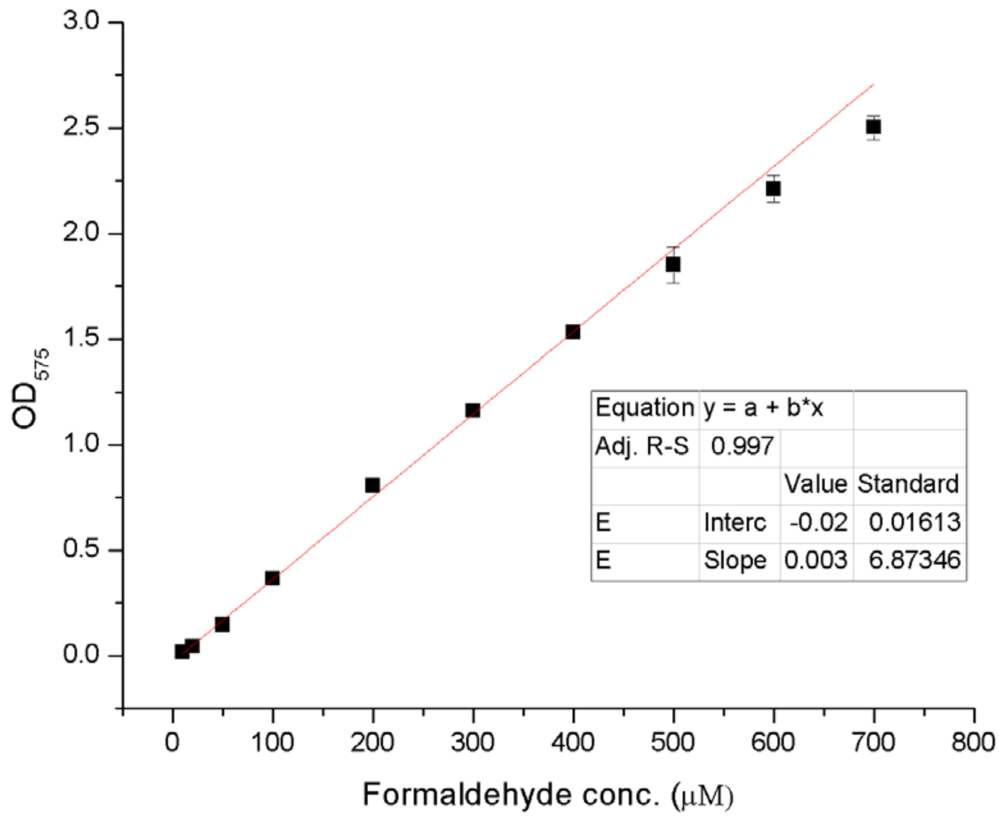


Figure 2.14. Standard curve obtained by the method of Topp and Knowles (1984) that was used to calculate the formaldehyde concentrations in the bacterial culture. Formaldehdye was diluted in MTYM medium to have a range of 10 to 700 µM.

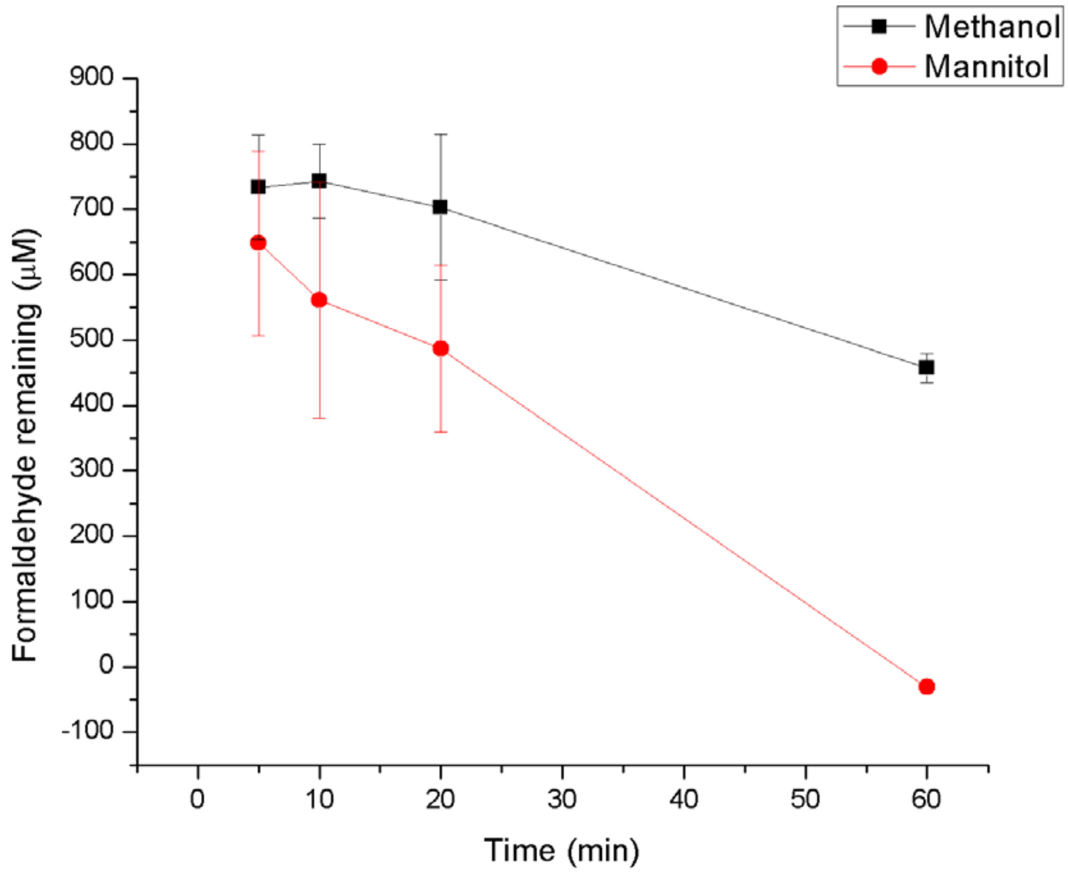


Figure 2.15. Remaining formaldehyde in the *B. methanolicus* MGA3 culture media. Formaldehyde (2 mM final) was added to *B. methanolicus* MGA3 growing in $MTYM_{MeOH}$ (Methanol) or $MTYM_{MAN}$ (Mannitol) at the OD_{600} of 0.8. One milliliter sample was used to measure the formaldehyde concentration in the culture media at 5, 10, 20 and 60 minutes post addition. (n=3)

CHAPTER 3

GENOME SEQUENCE of *Bacillus methanolicus* MGA3 and ELUCIDATION OF CARBON DISSIMILATION PATHWAY

Introduction

Advances in genome sequencing technologies greatly reduced the cost of whole genome sequencing. The time required to sequence a whole bacterial genome has also been reduced significantly since current sequencing technologies do not require cloning the DNA fragments into a bacterial host which is usually *E. coli*. These advances resulted in a geometric increase in the number of sequenced microorganisms (Binnewies et. al., 2006). The number of bacteria that have complete genome sequence has increased from 792 as of February 2009 (Tettelin and Feldblyum, 2009) to 2824 by February 2012 (GOLD, www.genomesonline.org). There are also 7931 ongoing bacterial genome sequence projects that are not complete yet. The presence of a vast amount of DNA sequences led to the birth of the genomics discipline.

The advances in sequencing also substantially widened the scope of metagenomic analysis of environmentally derived samples. In the pre-genomic era, it was required to isolate and culture a bacterium in the laboratory and run biochemical assays in order to obtain useful information about the metabolism of the bacterium. Having less than a half of 1% of 2-3 billion microbial species being identified (Fraser et. al., 2002), and only a very small proportion of these bacterial species being cultured (Handelsman, 2004), this limited our understanding of the microbial world and the metabolic diversity in the environment. However, advances in sequencing led researchers to overcome this difficulty. Qin et. al.

(2010) catalogued all microbial genes in the human gut, and constructed and compared the gut microbial flora between healthy and diseased individuals without isolating and culturing the individual bacteria. Similarly, Tyson et. al. (2004) revealed the metabolism and structure of a microbial community in the environment through reconstruction of microbial genomes from the environmental samples. Environmental genomes of SaragossaSea organisms were shotgun sequenced (Venter et. al., 2004) and revealed 1800 genomic species, 148 of which were unknown previously. This study also identified 1.2 million previously unknown genes. These studies show the power of sequencing; it would not be possible to isolate and culture all the 1800 microbial species and identify their unique genes.

Genome sequencing of obligate methanotrophic bacterium *Methylococcuscapsulatus* (Bath) (Ward et. al., 2004) led to the discovery of serine pathway genes in RuMP pathway methylotrophs which might have significant implications to the C1 metabolism of the bacterium. *M. capsulatus* (Bath) has a 3.3-Mb genome which encodes genes suggesting the ability to grow on sugars, to oxidize the hydrogen and sulfur and to grow under reduced oxygen tension. Another model microorganism that has been sequenced is the facultative methylotrophic bacterium *Methylobacteriumextorquens* AM1 (Vuilleumier et. al., 2009). It has a genome size of 6.88-Mb. Given both of these microorganisms are Gram-negative bacteria with high G+C content (63.6% and 68.7% for *M. capsulatus*(Bath) *and M. extorquens*AM1, respectively), it is necessary to sequence the genome of Gram-positive, thermotolerant, restrictive methylotroph*B. methanolicus* MGA3. Sequence of *B. methanolicus* MGA3 genome will serve as a reference to Gram-positive RuMP pathway methylotrophs.

Materials and Methods

Isolation of a plasmid from *B. methanolicus* PB1: The plasmid from *B. methanolicus* PB1 is isolated using Qiagen® Plasmid Midi Kit with the following modifications: 20 mg/ml lysozyme was added to Buffer P1 and incubated overnight at 37 °C and the mixture was centrifuged for 90 min at 14780g for the precipitation of plasmid with isopropanol at the precipitation step. The rest of the procedure is as written in the Qiagen manual.

Sequencing of the *B. methanolicus* PB1 plasmid: The plasmid from *B. methanolicus* PB1 was cloned into MCS of pUC18 after digestion with proper restriction enzymes. The clones were sequenced using M13 forward and reverse primers at DNA sequencing facility of University of Chicago. Primers were designed to sequence long inserts when needed.

Preparation of total DNA:*B. methanolicus*MGA3 was grown in MTYM media as described in Chapter 2 and 20 mL of cells were harvested by centrifugation (5 min x 10,000g) at the mid-exponential phase (OD_{600} of 1-1.5). The pellet was resuspended in 540 µl of lysis buffer (20 mMTris.Cl, 2 mM EDTA, 1.2% Triton X-100) with 20 mg/ml lysozyme and incubated at 37 °C for two hours. Total DNA was extracted with Qiagen® DNeasy® Blood and Tissue kit according to the manufacturer's instruction, albeit reagent volumes used were multiplied by 3 until filter binding. DNA was eluted in the buffer AE after extended incubation (up to 4 hours) of filter-bound DNA at room temperature. Eluted DNA was run on agarose gel (0.8% w/v) and quantified with a NanoDrop 1000.

Sequencing of *B. methanolicus*MGA3 total DNA and annotation: Sample preparation of *B. methanolicus* MGA3 total DNA was carried out at the North Carolina State University Genomics Sciences Laboratory (GSL). Total DNA of *B. methanolicus* MGA3 was shot-gun sequenced on the IlluminaGAIIx platform and raw sequences assembled *de novo* at GSL. Draft sequence of 117 contigs were analyzed and annotated using sequence visualization and annotation software Artemis (Trust Sanger Institute). Identified open reading frames (ORFs) were translated *in silico* and blasted against the protein database at NCBI.

Results

Sequence of plasmid from *B. methanolicus* PB1:Figure 3.1 presents the map of the plasmid isolated from *B. methanolicus* PB1. The sequences from the clones totaled to 20,187 bps. In compliance with the naming of pBM19 from *B. methanolicus* MGA3 (Figure 3.2), this plasmid is called pBM20. Like pBM19, pBM20 carries the RuMP pathway genes, *glpX, fba, tkt, pfk, rpe*along with *mdh.* The transposon *tnpI* is also present on the pBM20. The sequence of pBM20 presented in Appendix A.

Sequencing and assembly of *B. methanolicus* MGA3 genome: Theresults of the draft genome sequencing are summarized in Table 3.1 below. A total of 27,934,202 readings with 72 bp-long sequences yielded 2,011,262,544 bp total read length. There were 28,002 reads in 2,676,600 total sequence lengthin non-specific matches that were not mapped. *De novo* assembly of mapped reads resulted in a ~3.3 Mb genome from 117 contigs with an average coverage of 309 to 9091X. Based on the draft genome sequence obtained, *B. methanolicus* MGA3 genome has a GC content of 38.35%.

Table 3.1. Summary of *B. methanolicus* MGA3 genome sequencing

Contig count	117
Total read count	27,934,202
Mean read length	72.00
Total read length	2,011,262,544
Mean contig length	28,460
Total contig length	3,329,851
GC contents in %	38.35

Annotation: Annotation of the ORFs in *B. methanolicus* MGA3 draft genome is presented in Appendix B. The *B. methanolicus* MGA3 draft genome contains 2773 ORFs longer than 100 bp. Endogenous plasmid pBM19 sequence was assembled into Contig1. Previously reported methylotrophy genes (Brautaset et. al., 2004) were identified in the genome. Highlights from the genome sequence of *B. methanolicus* MGA3 are presented below.

Methanol dehydrogenase (*mdh*): Methanol dehydrogenase (MDH) is the first enzyme in the methanol utilization pathways of methylotrophs. Methanol is oxidized to formaldehyde by MDH. Methylotrophy in *B. methanolicus* was shown to be plasmid dependent since most of the RuMP genes and *mdh* were carried on the endogenous plasmid pBM19 (Brautaset et. al., 2004). Exceptions to these genes are hexulose phosphate synthase (*hps*) and phosphohexuloseisomerase (*phi)* which are located on the chromosome. Plasmid-borne *mdh* was shown to be the only gene responsible for the conversion of methanol to formaldehyde. Southern blots with plasmid born *mdh* as probe did not yield any hybridization on the chromosome. In addition, pBM19 cured *B. methanolicus* MGA3 was shown to lose the ability to grow on methanol. Genome sequence reveals two additional putative *mdh*genes in *B. methanolicus* MGA3; on contig27 and contig93 (ORFs 01267 and 02737). Both of these *mdh* genes are clearly different from the plasmid-borne *mdh*. The *mdh* on contig27 aligns with *Lysinibacillussphaericus*NAD-dependent methanol dehydrogenase. On the protein level, the *mdh* on contig27 has 70% identity with the *mdh* of *L. sphaericus*. The *mdh* on contig93 is similar to the *mdh* on contig27 with 96% identity on the protein level, and it aligns better with *L. fusiformis* ZC1. Both of these *mdh* genes are different from the previously reported methanol dehydrogenase (*mdh*) gene of *B. methanolicus* C1 (de Vries et. al., 1992). We also report the sequence of *mdh* from pBM20 of *B. methanolicus* PB1 for the first time. Comparison of *mdh* sequences is presented in Figure 3.3 and the relationship of the sequences are depicted in Figure 3.4.

Formaldehyde dehydrogenase (*faldh*): The next reaction in the dissimilation of methanol is the oxidation of formaldehyde to formate. Methanol was shown to be dissimilated through formaldehyde and formate in *B. methanolicus* MGA3 by ^{13}C NMR mass and isotope-ratio mass spectrometry (Pluschkell and Flickinger, 2002). Continuous cultures of *B. methanolicus* MGA3 were shown to accumulate formaldehyde and formate when spiked with [^{13}C] methanol at 50 °C (Pluschkell and Flickinger, 2002). Although this accumulation clearly demonstrates the conversion of methanol into formaldehyde and then formate, no *faldh* gene was reported from *B. methanolicus*. Attempts to demonstrate the presence of formaldehyde dehydrogenase by Southern blot using *faldh* from three *Bacilli* failed to provide the identity of the long sought genes (Smith, 2007 and Chapter 4).However, three ORFs in the genome of *B. methanolicus* MGA3 on contig19 (BMETHMGA300554c), contig20 (BMETHMGA300623) and contig36 (BMETMGA301733c) align with aldehyde dehydrogenases of *Geobacillus* sp. WCH70, *Geobacillusstearothermophilus,* and Bacillus sp. NRRL B-14911 with identities of 95%, 64%, and 91%, respectively.These genes are annotated as putative aldehyde dehydrogenases which might be the long soughtgenes. The FADH activities of the proteins encoded by these putative genes needs to be demonstrated biochemically.

THF- and THMPT-linkded enzymes: There is more than one possible route for conversion of formaldehyde into formate as mentioned in Chapter 1. *M. extorquens*AM1 oxidizes formaldehyde into formate mainly through THF- and THMPT-linked enzymes, although it also possesses a*faldh* gene (Crowther et. al., 2008). Putative genes encoding for the enzymes Methenyl-THF dehydrogenase and cyclohydrolase and Formate-THF ligase were found to be present in the genome of *B. methanolicus* MGA3. The presence of these enzymes suggests a possible alternative route for oxidation of C_1 compounds.

Formate dehydrogenase (*fdh)*: The last oxidation step in the dissimilation of methanol is the oxidation of formate into carbon dioxide. Oxidation of methanol to formate was known, as formate was the other intermediate shown to accumulate when *B. methanolicus* MGA3 was spiked with ^{13}C methanol (Pluschkell and Flickinger, 2002). Formate is thought to be oxidized to carbon dioxide by FDH since *B. methanolicus*MGA3 does not grow on formate. Although FDH activity was shown with crude cell extracts of *B. methanolicus* MGA3 (Smith, 2007), no *fdh* gene was reported. Two ORFs in the genome of *B. methanolicus*MGA3 encodes for putative FDHs. ORFs on contig32

(BMETHMGA301578c and BMETHMGA301579) encode two peptides; accessory protein and alpha chain of FDH, respectively. The alpha chain shares 81% identity with the *fdh* of *Bacillus halodurans* C-125. The accessory protein is on the complementary strand and transcribed in the opposite direction of the alpha chain. It shares 72% identity with formate dehydrogenase accessory protein of Bacillus sp. 2_A_57_CT2. The ORF on contig48 (BMETHMGA301931) encodes a putative *fdh* which is 82% identical to the *fdh* of *Geobacillus thermodenitrificans* NG80-2. As for putative formaldehyde dehydrogenases, the function of these *fdh* genes also needs to be verified biochemically.

Carbon monoxide dehydrogenase *(codh)*: Carbon monoxide can be used as a sole source of carbon and energy by microorganisms that employ the Wood-Ljungdahl pathway (Ragsdale, 2004). Carbon monoxide dehydrogenase (*codh*) plays a significant role in reduction of CO_2 and it has been extensively studied in anaerobic microorganisms (Ferry, 1999, Seravalli and Ragsdale, 2000, Ragsdale 2008). In *Moorella acetica* (formerly *Clostridium thermoaceticum*), it is found as a Carbon Monoxide Dehydrogenase/Acetyl-CoA Synthase (CODH/ACS) complex that reduces CO_2 to Acetyl-CoA by the Wood-Ljungdahl pathway (Seravalli and Ragsdale, 2000). The *codh* is widely distributed in Desulfuricants, Hydrogenogens, Acetogens and Methanogens (Oelgeschlager and Rother, 2008). The genome of *B. methanolicus* MGA3 also contains a cluster of ORFs (BMETHMGA301736-01742) that encodes for CODH protein subunits. The ORF BMETHMGA301738 encodes the 773 aa. large subunit of CODH that has 75% identity with the CODH of *Thermaerobacter subterraneus* DSM 13965. The ORFs BMETHMGA301736-01742 encode for CODH medium subunit, CODH small subunit, CODH large subunit, CODH subunit F, CODH subunit G, CODH subunit D and CODH subunit E, respectively. Considering the presence of THF- and THMPT linked enzymes together with CODH, possibility of anaerobic growth on CO_2 through Wood-Ljungdahl pathway deserves further attention and experimentation.

Conclusion

The sequencing of the genome of the *B. methanolicus* MGA3 is an important step in understanding the metabolism of this bacterium as a promising industrial candidate. Genome sequence analysis has revealed the presence of two more *mdh*s in the chromosome in addition to the plasmid borne gene. Putative genes of *aldh*s and *fdh*s were identified in the genome. In addition to the candidate genes with possible roles in dissimilation of methanol, *B. methanolicus* MGA3 genome sequence has also revealed the presence of putative THF- and THMPT linked enzymes that may function in dissimilation of methanol. Another important outcome of the genome sequence is the presence of putative *codh* in the *B. methanolicus* MGA3 genome which may offer versatility broader than previously thought (Brautaset et. al., 2007).

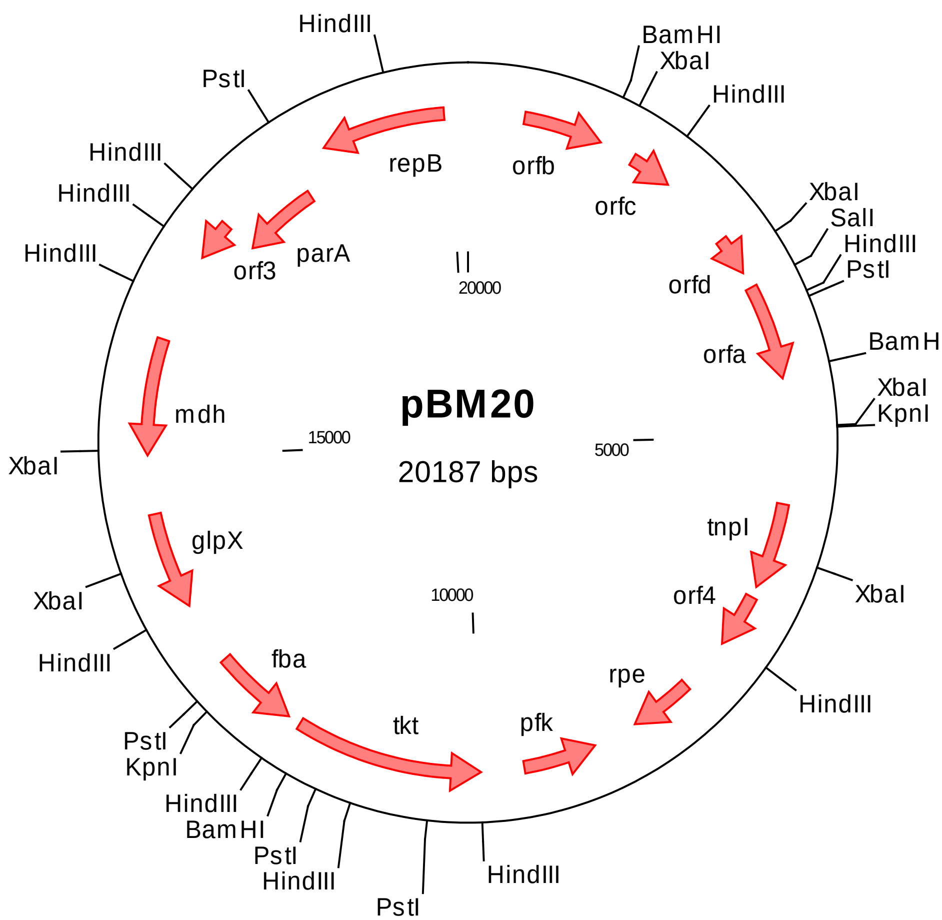


Figure 3.1. Plasmid from *B. methanolicus* PB1, pBM20.

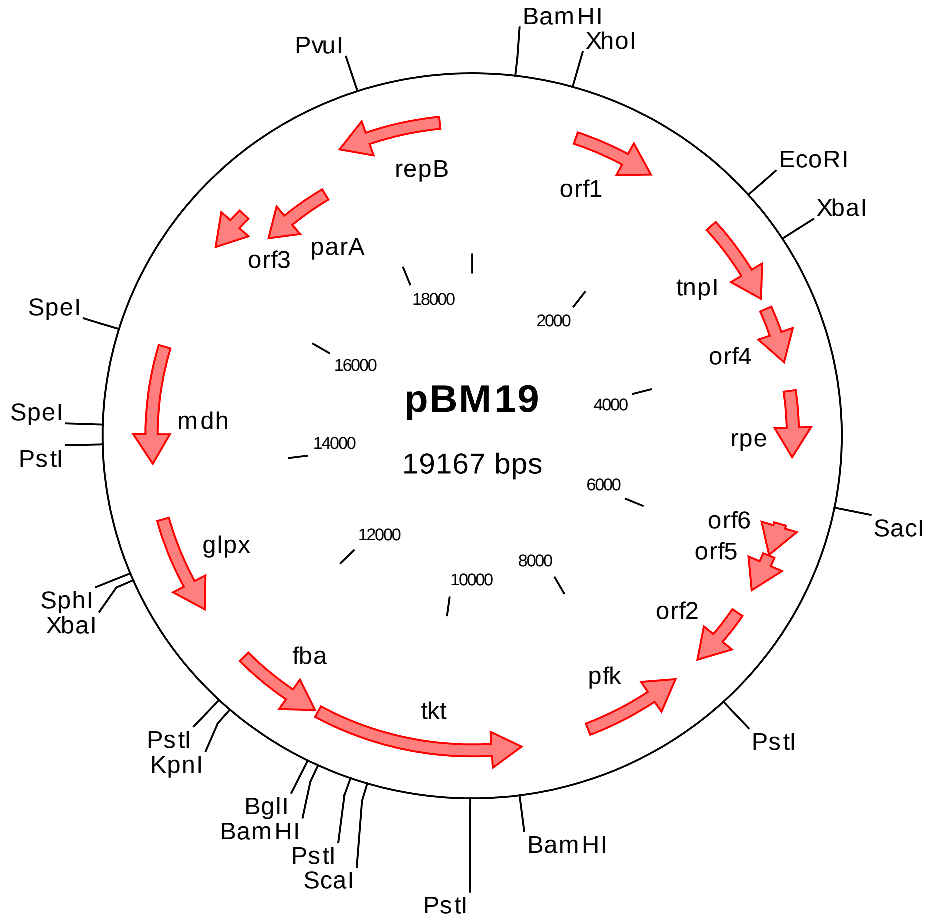


Figure 3. 2. Plasmid from *B. methanolicus* MGA3, pBM19 (Brautaset et. al., 2004).

```
B.methC1        ----MTNFFIPPASVIG--AVKEVGTRLKQIG--KALIVTDAFLHSTG--EEVAKNIREA 50
B.methPB1       --MTQRNFFIPPASVIGRGAVKEVGTRLKQIGATKALIVTDAFLHGTGLSEEVAKNIREA 58
B.methMGA3      ---MTTNFFIPPASVIGRGAVKEVGTRLKQIGAKKALIVTDAFLHSTGLSEEVAKNIREA 57
B.meth27        MKNTQSAFYMPSVNLFGAGSVNEVGTRLAGLGVKKALLVTDAGLHSLGLSEKIAGIIREA 60
B.meth93        MTNTQSAFFMPSVNLFGAGSVNEVGTRLADLGVKKALLVTDAGLHGLGLSEKISSIIRAA 60
                       *::*...::*  :*:******  :*  ***:**** **. *  *:::  ** *

B.methC1        GLD--VFPKAQPDPADTQVH--VDVFKQENCDALVS--GGSSHDTAKAIGLV--NGGRIN 102
B.methPB1       GLDAVIFPKAQPDPADTQVHEGVDIFKQEKCDALVSIGGGSSHDTAKAIGLVAANGGRIN 118
B.methMGA3      GVDVAIFPKAQPDPADTQVHEGVDVFKQENCDSLVSIGGGSSHDTAKAIGLVAANGGRIN 117
B.meth27        GVEVAIFPKAEPNPTDKNVAEGLEAYNAENCDSIVTLGGGSSHDAGKAIALVAANGGTIH 120
B.meth93        GVEVSIFPKAEPNPTDKNVAEGLEAYNAENCDSIVTLGGGSSHDAGKAIALVAANGGKIH 120
                *::  :****:*:*:*.:*   :: :: *:**::*:  ******:.***.**  *** *:

B.methC1        DYQGVNSV--PVVPVVAITTTAGT--ETTSLAVITDSAR--KMPVIDEKITPTVAI--PE 154
B.methPB1       DYQGVNSVEKPVVPVVAITTTAGTGSETTSLAVITDSARKVKMPVIDEKITPTVAIVDPE 178
B.methMGA3      DYQGVNSVEKPVVPVVAITTTAGTGSETTSLAVITDSARKVKMPVIDEKITPTVAIVDPE 177
B.meth27        DYEGVDVSKKPMVPLIAINTTAGTGSELTKFTIITDTERKVKMAIVDKHVTPTLSINDPE 180
B.meth93        DYEGVDVSKEPMVPLIAINTTAGTGSELTKFTIITDTERKVKMAIVDKHVTPTLSINDPE 180
                **:**:    *:**::**.*****  * *.:::***: *  **.::*:::***::*  **

B.methC1        LMVKKPAGLTIA--MDALSHAIEAYVAK--TPVTDAFAIQAMKL--EYLPKAVANGEDIE 208
B.methPB1       LMVKKPAGLTIATGMDALSHAIEAYVAKRATPVTDAFAIQAMKLINEYLPRAVANGEDIE 238
B.methMGA3      LMVKKPAGLTIATGMDALSHAIEAYVAKGATPVTDAFAIQAMKLINEYLPKAVANGEDIE 237
B.meth27        LMVGMPPSLTAATGLDALTHAIEAYVSTGATPITDALAIQAIKIISKYLPRAVANGKDIE 240
B.meth93        LMVGMPPSLTAATGLDALTHAIEAYVSTGATPITDALAIQAIKIISKYLPRAVANGKDIE 240
                ***  *..** *  :***:*******:.  **:***:****:*:  :***:*****:***

B.methC1        --EAMAYAQYMAGVAF--GGLGLVHSISHQVG--YKLQHGICNSVNMP--CAFNLIAKTE 260
B.methPB1       AREAMAYAQYMAGVAFNNGGLGLVHSISHQVGGVYKLQHGICNSVNMPHVCQFNLIARTE 298
B.methMGA3      AREKMAYAQYMAGVAFNNGGLGLVHSISHQVGGVYKLQHGICNSVNMPHVCAFNLIAKTE 297
B.meth27        AREQMAFAQSLAGMAFNNAGLGYVHAIAHQLGGFYNFPHGVCNAILLPHVCRFNLISKVE 300
B.meth93        AREQMAFAQSLAGMAFNNAGLGYVHAIAHQLGGFYNFPHGVCNAVLLPYVCRFNLISKVE 300
                  * **:** :**:**  .*** **:*:**:*  *:: **:**:: :*  * ****::.*

B.methC1        RFAH--ELLGENVSGLSTA--AERAIVALERYNKNF--PSGYAEMGVKEE--DILAKNAF 312
B.methPB1       RFAHIAELLGENVSGLSTASAAERAIVALQRYNKNFGIPSGYAEMGVKEEDIELLANNAY 358
B.methMGA3      RFAHIAELLGENVAGLSTAAAAERAIVALERINKSFGIPSGYAEMGVKEEDIELLAKNAY 357
B.meth27        RYAEIAAFLGENVDGLSTYEAAEKAIKAIERMARDLNIPKGFKELGAKEEDIETLAKNAM 360
B.meth93        RYAEIAAFLGENVDGLSTYDAAEKAIKAIERMAKDLNIPKGFKELGAKEEDIETLAKNAM 360
                *:*.   :***** ****   **:** *::*  :.:  *.*: *:*.***  : **:**

B.methC1        EDVCTQSN--VATVQDIAQIIKNAGALAEIAEGVVTGIANGHIAGEPLRKSIGGAKSKDG 370
B.methPB1       QDVCTLDNPRVPTVQDIAQIIKNAL----------------------------------- 383
B.methMGA3      EDVCTQSNPRVPTVQDIAQIIKNAM----------------------------------- 382
B.meth27        NDACALTNPRKPKLEEVIQIIKNAM----------------------------------- 385
B.meth93        KDACALTNPRKPKLEEVIQIIKNAM----------------------------------- 385
                :*.*:  *   ..:::: ******

B.methC1        ANRNVVAAILR 381
B.methPB1       -----------
B.methMGA3      -----------
B.meth27        -----------
B.meth93        -----------
```

Figure 3.3. Alignment of methanol dehydrogenase sequences from *B. methanolicus* strains. B.methC1; *B. methanolicus* C1, B.methPB1; *B. methanolicus* PB1, B.methMGA3; *B. methanolicus* MGA3, B.meth27; *B. methanolicus* MGA3 ORF01267, B.meth93; *B. methanolicus* MGA3 ORF02737. "*"; identical in all sequences, ":"; conserved substitutions are observed, "."; semi-conserved substitutions are observed.

Figure 3.4. Relatedness of methanol dehydrogenase sequences from *B. methanolicus* strains by Neighbour-Joining method (Thompson et. al., 1994). B.methC1; *B. methanolicus* C1, B.methPB1; *B. methanolicus* PB1, B.methMGA3; *B. methanolicus* MGA3, B.meth27; *B. methanolicus* MGA3 ORF 01267, B.meth93; *B. methanolicus* MGA3 (ORF 02737)

CHAPTER 4

ELUCIDATION OF C_1 ASSIMILATION AND DISSIMILATION PATHWAYS IN *B. methanolicus*

Introduction

Methanol assimilation pathway: The RuMP pathway was briefly described in Chapter 1. *B. methanolicus* employs the FBPA/SBPase variant of the RuMP pathway. There are two enzymes that function in the fixation part. The first enzyme of the RuMP pathway is the hexulose phosphate synthase (HPS). This enzyme catalyzes the synthesis of hexulose-6-phosphate from ribulose-5-phosphate and formaldehyde produced by the oxidation of methanol by methanol dehydrogenase. Three molecules of formaldehyde are condensed with ribulose-5-phosphate to yield three molecules of hexulose-6-phosphate. The second reaction in the fixation part is the isomerisation of hexulose-6-phosphate to fructose 6-phosphate (F6P) by 6-phospho-3-hexuloisomerase (PHI). The ORFs BMETHMGA300426 and BMETHMGA300427 located on the chromosome annotated as *hps* and *phi* encode for the enzymes HPS and PHI respectively. The first reaction in the cleavage part is the conversion of the F6P to fructose-1,6-bisphosphate (FBP) by phosphofructokinase (PFK). FBP is then cleaved to glyceraldehyde 3-phosphate (GAP) and dihydroxyacetone phosphate (DHAP) by fructose-1,6-bisphosphate aldolase. GAP enters into the regeneration part along with two FBPs where three ribulose 5-phosphate molecules are regenerated while three-carbon DHAP is the net gain of the RuMP pathway. *B. methanolicus* MGA3 is known to carry both genes encoding for PFK and FBPA on pBM19 (Brautaset et al., 2004). There are three more RuMP pathway genes that are located on pBM19, namely transketolase (*tkt*), SBPase (*glpX*), and ribulose 5-phosphate 3-epimerase (*rpe*) (Brautaset et. al., 2004).

Methanol dissimilation pathway: The methanol dissimilation pathway, oxidation of methanol to carbon dioxide through formaldehyde and formate, is not as well characterized in *B. methanolicus* as the RuMP pathway. Although the formation and accumulation of formaldehyde and formate were demonstrated with ^{13}C NMR studies by Pluschkell and Flickinger (2002) attempts to identify possible formaldehyde and formate dehydrogenases from *B. methanolicus*by PCR failed (Smith, 2007).

This chapter reports the identification of formaldehyde and formate dehydrogenases in *B. methanolicus* strain PB1 by PCR and Southern Blot, and also proposes the possibility of an alternative dissimilation pathway through the presence of THF- and THMPT-linked enzymes. This chapter also reports the presence of additional copies of RuMP pathway genes on the chromosome.

Materials and Methods

Cloning *B. methanolicusfaldh*:Total DNA was isolated from *B. methanolicus* and *B. subtilis*, *Geobacillus kaustophilus* and *Bacillus licheniformis* as described in Chapter 3. Primers with indicated restriction sites at the end (Table 4.1) were designed specific to the putative aldehyde dehydrogenases (*aldh*) that have high homology to known formaldehyde dehydrogenases (*faldh)*. The *yycR* gene of *B. subtilis*, *ywdH* gene of *B. licheniformis* and *G. kaustophilus* ORF between 2797314-2798810 bases were selected as candidate *faldh* genes. These primers were used to amplify candidate *faldh* genes by PCR using the total DNA of *B. methanolicus* strains MGA3 and PB1 with the following parameters: 5 min initial denaturation at 94 °C then 25 cycles of 1 min annealing at 52 °C, 1 min extension at 72 °C, and 30 sec denaturation at 94 °C. A final extension at 72 °C was added at the end of the cycle. The PCR reactions contained 500 nM forward and reverse primers, 1.5 mM $MgCl_2$, 500 µM of dNTP mix, and 50-200 ng of template DNA. The PCR mixture was then run on a 0.8 % agarose gel and visualized with Kodak Gel Documentation. The PCR fragment then was digested with respective restriction enzymes and ligated into pET16-b or pET17-b. *E. coli* DH5α was transformed with the plasmid as described by Maniatis et al., (1982), and the colonies were screened for the insert using T7 forward and reverse primers with the PCR parameters as described above.

Cloning of *B. methanolicus fdh*: The same approach was used for the cloning of *fdh* genes of *B. methanolicus* using the primers designed for putative *fdh* (*yrhE*) genes of *B. subtilis*, *B. licheniformis* and *G. kausotphilus* ORF located in between bases 498622-501585 using the primers listed in Table 4.1.

Southern Blot: Southern blotting was employed to demonstrate the presence of putative *faldh* and *fdh* genes in *B. methanolicus* strains. The primers in Table 4.1 were used for the synthesis of Digoxigenin (DIG)-labeled DNA probes by Roche's PCR DIG Probe Synthesis Kit (Cat# 11 636 090 910) as described in the manufacturer's manual using the total DNA from the respective *Bacilli* species as template. A DIG-labeled DNA probe was also synthesized for the citrate synthase (*citY*) gene of *B. methanolicus* MGA3 as a control using Bm_CitY_Fw and Bm_CitY_Rv primers in Table 4.1. Incorporation of DIG-labeled nucleotides into the PCR amplicon was checked with gel electrophoresis as DIG-labeled DNA moves slower than its non-labeled counterpart. The total DNA from *B. methanolicus* strains were digested completely using restriction enzyme NdeI. Different amounts of DNA were blotted to find the required amount for the hybridization and 12 µg was found to give good hybridization. Twelve µg of digested DNA was loaded into an agarose gel and run until a good separation was achieved. Then the DNA was transferred onto a nylon membrane by capillary transfer as described in the Roche's application manual. Prehybridization, hybridization, washing and detection were carried out according to the instructions in the manual. Hybridization was carried out at 55 °C in 50 ml falcon tubes in HYBAID incubator. The membrane was washed and blocked using the Roche's DigWash and Block Buffer Set (Cat# 11 585 762 001). Anti-Digoxigenin-AP, Fab fragments (Cat# 11 093 274 910) along with the NBT/BCIP tablets, were used for the detection and visualization of the hybridization.

Results

RuMP pathway genes: *B. methanolicus* is known to employ the RuMP pathway for the assimilation of C_1-carbon into biomass, and it was shown to have a plasmid-dependent methylotrophy (Brautaset et al., 2004) since the entire RuMP pathway genes except the *hps* and *phi* were carried on a native plasmid pBM19. Here we describe additional copies of RuMP pathway genes on the chromosome: The *B. methanolicus* MGA3 ORF BMETHMGA301123 encodes a ribulose phosphate epimerase enzyme which has 80% identity to *Bacillus* sp. 2_A_57_CT2. It also shares a 74% identity with the plasmid borne *rpe*, the ORF BMETHMGA300014c. The *pfk* gene also has an additional copy on the chromosome; BMETHMGA300592. It is 81% identical to *pfk* of *Bacillus* sp. NRRL B-14911and 61% identical to plasmid-borne *pfk*. The chromosomal copy of the *tkt* gene is located on the ORF BMETHMGA300446c. The chromosomal *tkt* has 84% identity to *Bacillus* sp. 2_A_57_CT2 and shares 76% identity with *tkt* carried on the pBM19. *B. methanolicus* MGA3 has a chromosomal copy of *fba* on ORF BMETHMGA302111 which has 88% identity to *fba* of *B. licheniformis* ATCC 14580. The ORF BMETHMGA302111 shares 74% identity with the previously reported *fba* on pBM19 (Brautaset et al., 2004). The chromosomal copy of SBPase (*glpX*) is located down stream of the *fba*, on ORF BMETHMGA302113. It has 87% identity to the *glpX* gene of *B. licheniformis* ATCC14580 and a comparatively lower identity to the plasmid borne *glpX*, 54%. In summary, all of the RuMP genes carried on the pBM19 have chromosomal counterparts although they are not identical.

Cloning *faldh* of *B. methanolicus* PB1: The primers in Table 4.1 were used to target *faldh* genes of *B. methanolicus* strains MGA3 and PB1 by PCR. The annealing temperature of the PCR was varied to optimize the amplification and the best temperature was found to be 52 °C. As can be seen in the picture of the gel electrophoresis (Figure 4.1), only primers designed for *B. licheniformis ywdH* produced amplification. This fragment was digested with NdeI and BamHI sites incorporated into the primers and cloned into the respective site of pET16-b. The insert was sequenced with T7 forward, T7 reverse. The partial sequence obtained shares 97% identity with the *ywdH* gene of *B. licheniformis* ATCC14580.

Southern Blot: Attempts to clone prospective *faldh* and *fdh* genes from *B. methanolicus* were unsuccessful with *B. licheniformis yydH* primers except for the putative *faldh* gene from strain PB1. Therefore, I aimed to show the presence of these genes in *B. methanolicus* by Southern blotting. The restriction enzyme NdeI produces a good smear as can be seen in Figure 4.2. The *B. methanolicus citY* gene was used as a control. Figure 4.3 clearly shows the efficiency of hybridization with different amounts of DNA as template. It can be seen from the Figure 4.4 that there is similarity of DNA between strains MGA3 and PB1, and *B. methanolicus* PB1 contains three copies of the citrate synthase gene. Figures 4.5-4.7 displays the hybridization of probes from genes of three *Bacilli* with *B. methanolicus* PB1 DNA. As can be seen, *B. methanolicus* PB1 DNA hybridizes with the *B. licheniformisywdH* gene but the *B. subtilis* or *G. kaustophilus* probes did not produce any hybridization with either of the strains. As for the *fdh* (Figures 4.8-4.10), the probes from *B. subtilis*, *B. licheniformis* and *G. kaustophilus* all hybridized with the *B. methanolicus* PB1 DNA although the hybridization of *B. subtilis* was weaker than the others. The absence of hybridization with the *B. methanolicus* MGA3 DNA is most probably due to DNA structure of the strain, since putative *fdh* of strains MGA3 and PB1 are 90% identical. Similarly, putative *fdh* of *G. kaustophilus* is 53% identical to *fdh* of PB1, and 54% to the *fdh* of strain MGA3. However, it only hybridizes with strain PB1 DNA, which suggests difference in the DNA structure rather than DNA sequence between the strains MGA3 and PB1.The relative places of the hybridization bands on the membrane suggest that *B. methanolicus* PB1 carries more than one copy of the *fdh* gene.

None of the tested probes produced hybridization with *B. methanolicus* MGA3 while they hybridized with strain PB1 DNA. This might be either because of the structure of the DNA or the difference in the base composition of the target genes. To test this, probes of *B. subtilis citZ* gene were hybridized against both strains. Although *B. methanolicus* MGA3 has a citrate synthase II gene, *citY* (Brautaset et. al., 2003),the *B. subtilis citZ* probe hybridized only with PB1 DNA as it can be seen in Figure 4.11. It is also noteworthy that putative *fdh* genes and *aldh* genes, possibly *faldh* were identified in the genome of *B. methanolicus* MGA3 (below). This suggests that the base composition of the genes, *citY* gene in this case, in strain MGA3 is different than strain PB1.

Methanol dissimilation pathways of *B. methanolicus*: The presence of putative *faldh* and *fdh* genes of *B. methanolicus* in addition to chromosomal *mdh* genes were explained in Chapter 3. However, these three genes are not the only possible route for the dissimilation of methanol. The presence of THF- and THMPT-linked pathway genes also suggests alternative routes for carbon dissimilation. Figure 4.12 outlines the possible routes methanol carbon may follow to carbon dioxide. The first reaction of the THF-linked pathway, the condensation of formaldehyde with THF, occurs non-enzymatically. The *B. methanolicus* MGA3 genome possesses a putative methenyl-THF cyclohydrolase & dehydrogenase gene (ORF 00712) that is 76 % identical to the *Bacillus sp.* 1NLA3E. This is a bifunctional enzyme that can catalyze the conversion of methylene-THF to methenyl-THF first and then to formyl-THF. It is possible for the THF- and THMPT-linked enzymes to cross-catalyze the substrates as in the example of the MtdA enzyme of *M. extorquens* AM1 (Hagemeier et. al. 2000), and these enzymes may function in detoxification of formaldehyde (Marx et. al., 2003). The next reaction of the pathway is the conversion of the formyl-THF into formate. This reaction is catalyzed by formyl-THF cyclo-ligase (ORF BEMTHMGA300665). This putative enzyme shares 58% identity with the formyl-THF cyclo-ligase of *Bacillus* sp. 2_A_57_CT2. The last step in the dissimilation is the oxidation of formate into carbon dioxide by formate dehydrogenase. As discussed in Chapter 3, *B. methanolicus* MGA3 possesses two putative formate dehydrogenases.

Table 4.1. List of primers used for cloning of respective genes and synthesis of DIG-labeled probes. Respective restriction sites are underlined.

Microorganism	Gene	Primer	Sequence
B. subtilis	*yycR*	Bs_faldh_Fw	5'-ATG GTA CTC GAG ATG GGA GGA ATG GCG TTG ACA G-3'(XhoI)
B. subtilis	yycR	Bs_faldh_Rv	5'-ATG GTA GGA TCC TTA CTT CAA TGT CCC ATG-3' (BamHI)
B. subtilis	yrhE	Bs_fdh_Fw	5'-ATG GTA CTC GAG ATG AGA TCT TTT GCC ACA CAA C- 3'(XhoI)
B. subtilis	yrhE	Bs_fdh_Rv	5'-ATG GTA GGA TCC TCA ATC TGT ACA CGG CCT G-3' (BamHI)
B. licheniformis	ywdH	Bl_faldh_Fw	5'- ATG GTA CAT ATG GCA GAA CGA ACA ACC TCC-3' (NdeI)
B. licheniformis	ywdH	Bl_faldh_Rv	5'-ATG GTA GGA TCC CTA CCT GAG AAG TTT TTT G-3' (BamHI)
B. licheniformis	yrhE	Bl_fdh_Fw	5'- ATG GTA CAT ATG GAC GGA AAG CCA ATC AG-3' (NdeI)
B. licheniformis	yrhE	Bl_fdh_Rv	5'- ATG GTA GGA TCC TTA GTC TGT TAA ATG AAC G-3' (BamHI)
G. kaustophilus	2797314-	Gk_faldh_Fw	5'- ATG GTA CAT ATG ACA ATG ATT TCG GCG-3' (NdeI)
G. kaustophilus	2797314-	Gk_faldh_Rv	5'- ATG GTA GGA TCC TTA AAA CAA CGT CAT CGG-3' (BamHI)
G. kaustophilus	498622-	Gk_fdh_Fw	5'-ATG GTA CAT ATG GAA GCG CAA ATG GTG-3'(NdeI)
G. kaustophilus	498622-	Gk_fdh_Rv	5'-ATG GTA GGA TCC TTA GCC ACC GAT CGT CAC-3' (BamHI)
B. subtilis	citZ	Bs_CitZ-Fw	5'-ATG ACA GCG ACA CGC GG-3'
B. subtilis	citZ	Bs_CitZ-Rv	5'-TTA GGC TCT TTC TTC AAT CG-3'
B. methanolicus		Bm_citY_Fw	5'-ATG ACA GTT ACA CGT GGT C-3'
B. methanolicus		Bm_citY_Rv	5'-TTA TCC TCT TTG CTC AAT TGG-3'

Wood-Ljungdahl pathway and Carbon monoxide dehydrogenase: The anaerobic bacterium *Clostridium ljungdahlii* uses the Wood-Ljungdahl pathway to fix carbon dioxide. The enzymology of the pathway is discussed in detail by Ragsdale (2008). The THF-linked enzymes mentioned above also take part in the fixation of carbon from carbon dioxide (Ragsdale, 2008). With the addition of methylene tetrahydrofolate reductase and carbon monoxide dehydrogenase enzymes, all enzymes of the eastern branch of the Wood-

Ljungdahl pathway can be located in the genome of *B. methanolicus* MGA3 (Figure 4.13). Briefly, carbon dioxide is reduced to formate by formate dehydrogenase. Then, it is condensed with THF to form formyl-THF by formate-THF ligase (ORF BMETHMGA300091). The putative *B. methanolicus* MGA3 formate-THF ligase is a 562 aa protein with 82% identity to the gene of *Bacillus* sp. 2_A_57_CT2. The next two reactions, conversion of the formyl-THF to 5,10-methenyl-THF and then to 5,10-methylene-THF are catalyzed by the methenyl-THF cyclohydrolase&dehydrogenase enzyme described above. Then, methylene-THF reductase (*metF*) converts 5,10-methylene-THF into 5-methyl-THF. *B. methanolicus* MGA3 possesses two copies of *metF*, ORFs BMETHMGA300824c and BMETHMGA301597c. The former 310 aa-long *metF* has 97% identity to the *metF* gene of *Geobacillus thermoglucosidans* while the latter 621 aa-long *metF* shares 77% identity with the *metF* gene of *Bacillus* sp. 2_A_57_CT2. Then, the methyl group of methyl-THF is condensed with carbon dioxide by carbon monoxide dehydrogenase (CODH) to produce acetyl-CoA. CODH enzymes may be bifunctional and also catalyze the reaction of Acetyl-CoA synthase enzyme (ACS). *B. methanolicus* MGA3 has a multimeric carbon monoxide dehydrogenase. The *codh* operon consists of seven ORFs, BMETHMGA301736-42. The large subunit of the *B. methanolicus* MGA3 *codh* is 75% identical to *Thermaerobacter subterraneus codh* large subunit. It should also be noted here that *T. subterraneuscodh* encodes an aerobic-type CODH which catalyzes the oxidation of carbon monoxide to carbon dioxide.

Conclusion

The genomic data revealed the presence of additional chromosomal copies of all RuMP pathway genes except the *hps* and *phi*. Being the first two reactions of the RuMP pathway, it is clear that the rate of down stream reactions will depend on the rates of these two enzymes and upregulation of these enzymes will result in an increased rate of the overall RuMP pathway (Brautaset 2004). The southern blot results demonstrate that the DNA composition of *B. methanolicus* strain MGA3 is different from the strain PB1. Combined with the southern blot results, genomic data suggest the presence of the *faldh* and *fdh* genes in both MGA3 and PB1 strains. For the first time, we propose putative *faldh* and *fdh* genes in *B. methanolicus* MGA3. In addition to *faldh* and *fdh* genes, the presence of the THF-linked enzymes implies an alternative route for oxidation of the formaldehyde to formate as depicted in Figure 4.12

Bacillithiol is a newly described antioxidant compound in *Bacillus*species (Newton et. al., 2009) that functions similar to mycothiol. *B. methanolicus* carries some of the bacillithiol synthesis genes, like*Bsh*A(ORF 2041) which suggests the synthesis of bacillithiol. However, no *faldh* enzyme was described to be bacillithiol dependent.The proposed formaldehyde dissimilation pathway does not change in the presence of bacillithiol in *B. methanolicus*.

Figure 4.13 displays a possible Wood-Ljungdahl pathway in *B. methanolicus* MGA3. With the inclusion of methenyl-THF reductase (BMETHMGA300824c and 01597c), the Wood-Ljungdahl pathway is completed. The presence of this pathway is rather surprising for an aerobic methylotrophic microorganism. As stated by Cordwell (1999), the presence of a gene sequence does not directly imply that the enzyme is functional. The gene might be a remnant nonactive gene through the evolutionary history. The functionality of these genes needs to be confirmed biochemically.

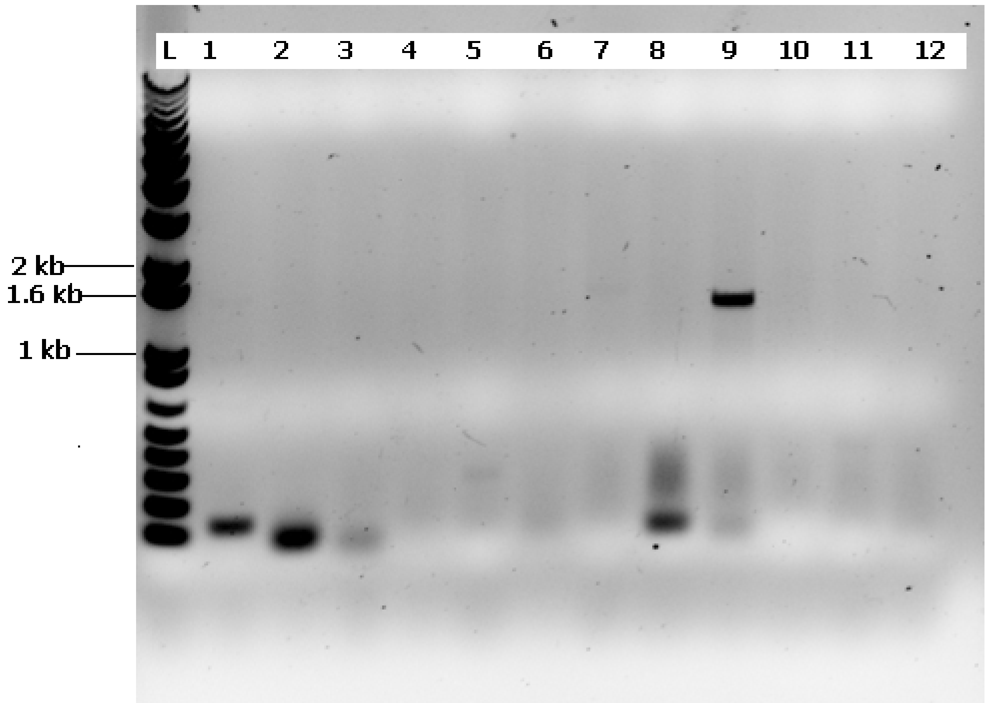


Figure 4.1. PCR amplification of *faldh*from*B. methanolicus* MGA (wells 1-6) and PB1 strains (wells 7-12) with *G. kaustophilus* putative *faldh* (1 and 7) and *fdh* (2 and 8), *B. licheniformisywdH* (3 and 9) and *yrhE* (4 and 10), *B. subtilis yycR* (5 and 11) and *yrhE* (6 and 12) primers. *B. licheniformisywdH* primers produced a similar length (1500 bp) amplicon from strain PB1.

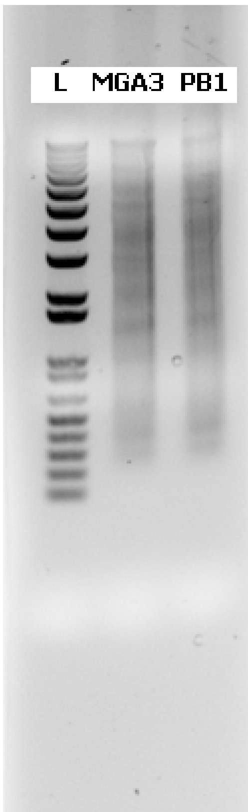


Figure 4.2. Digest of *B. methanolicus* strains MGA3 and PB1 with restriction enzyme NdeI. L; DNA ladder, MGA3: total DNA from strain MGA3, PB1; total DNA from strain PB1.

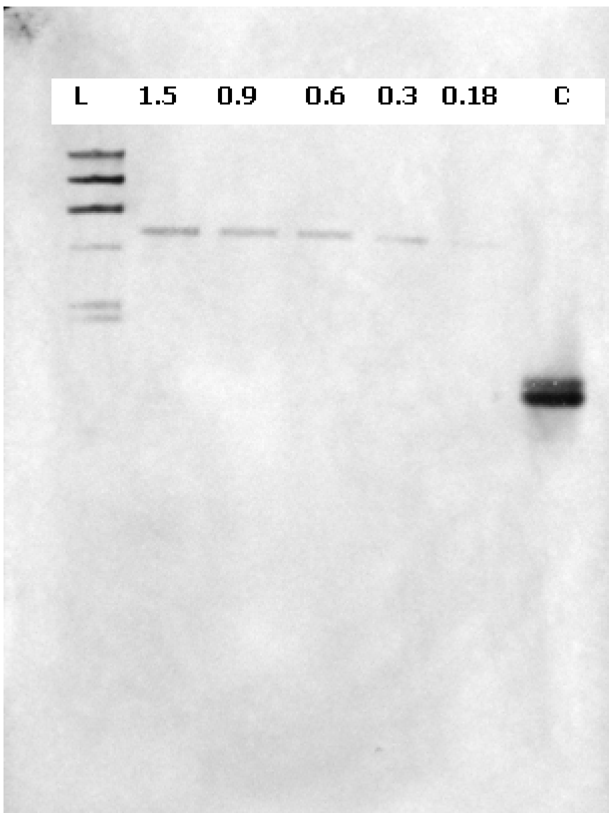


Figure 4.3. Southern blot control experiment with *B. methanolicus* MGA3 *citY* gene as probe and *B. methanolicus* MGA3 DNA digested with NdeI as target. L; Dig-labeled ladder, 1.5, 0.9, 0.6, 0.3 and 0.18; amount of DNA loaded to the gel before blotting to the membrane. C; control with *B. methanolicus* MGA3 *citY* gene, amplified by PCR.

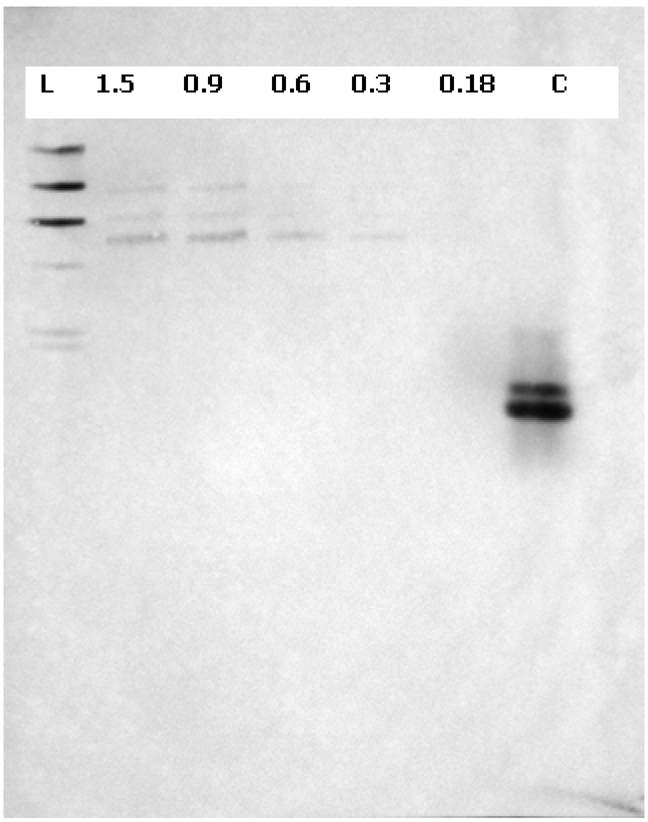


Figure 4.4. Hybridization of *B. methanolicus* MGA3 *citY* gene as probe with *B. methanolicus* PB1 DNA digested with NdeI as target. L; Dig-labeled ladder, 1.5, 0.9, 0.6, 0.3 and 0.18; amount of DNA loaded to the gel before blotting to the membrane (µg). C; control with *B. methanolicus* MGA3 *citY* gene, amplified by PCR.

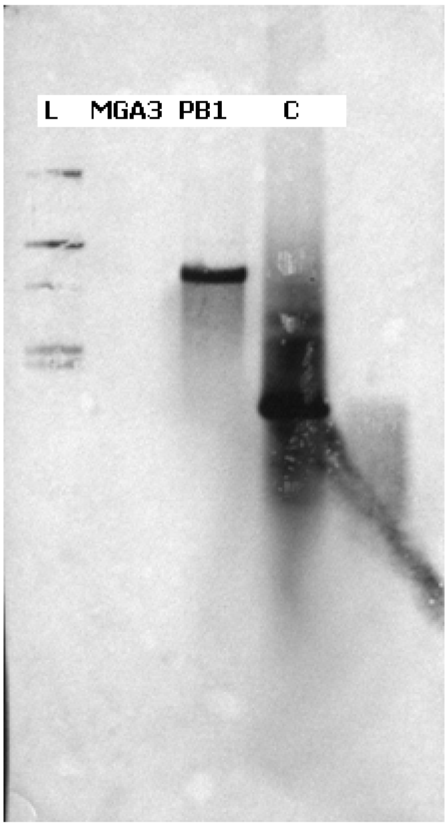


Figure 4.5. Hybridization of *B. licheniformi sywdH* gene as probe against*B. Methanolicus* MGA3 and PB1. L; Dig-labeled DNA ladder, MGA3; *B. methanolicus* MGA3, PB1; *B. methanolicus* PB1, C; control (PCR amplicon of *B. licheniformisywdH* gene).

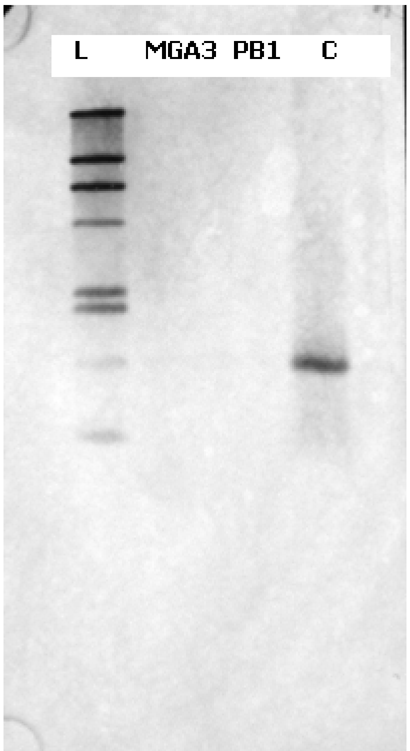


Figure 4.6. Hybridization of *B. subtilis ywdH* gene as probe against *B. methanolicus* MGA3 and PB1. L; Dig-labeled DNA ladder, MGA3; *B. methanolicus* MGA3, PB1; *B. methanolicus* PB1, C; control (PCR amplicon of *B. subtilis ywdH* gene).

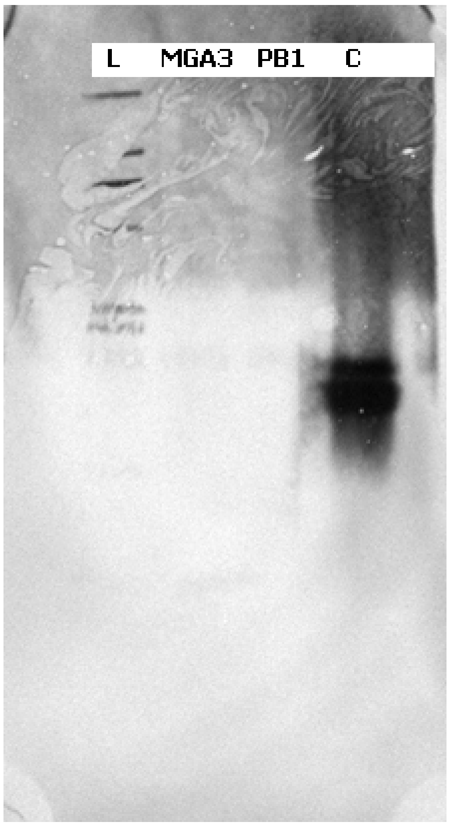


Figure 4.7. Hybridization of *G. kaustophilus* putative *faldh* gene as probe against *B. methanolicus* MGA3 and PB1. L; Dig-labeled DNA ladder, MGA3; *B. methanolicus* MGA3, PB1; *B. methanolicus* PB1, C; control (PCR amplicon of *G. kaustophilus*putative *faldh* gene).

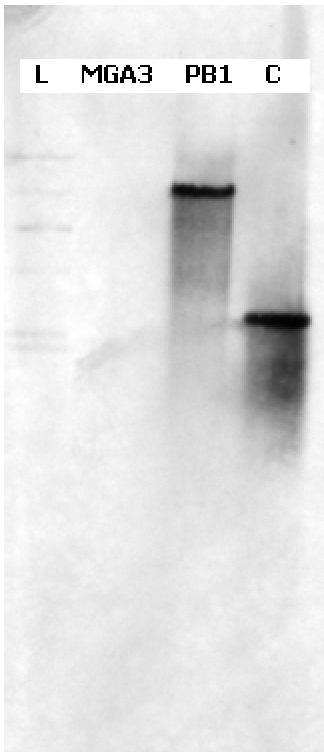


Figure 4.8. Hybridization of *B. licheniformis yrhE* gene as probe against *B. methanolicus* MGA3 and PB1. L; Dig-labeled DNA ladder, MGA3; *B. methanolicus* MGA3, PB1; *B. methanolicus* PB1, C; control (PCR amplicon of *B. licheniformisyrhE* gene).

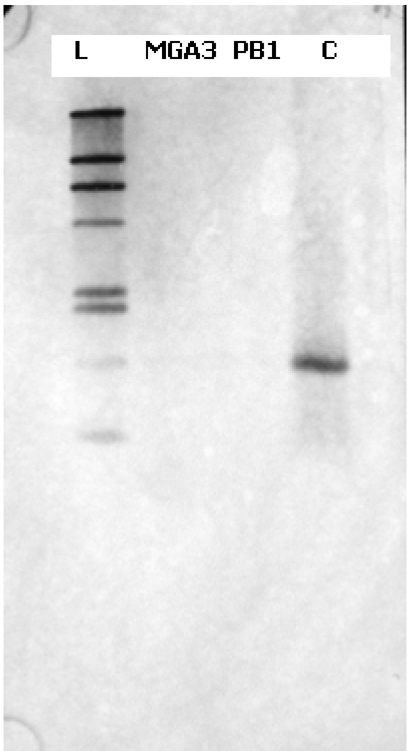


Figure 4.9. Hybridization of *B. subtilis yrhE* gene as probe against *B. methanolicus* MGA3 and PB1. L; Dig-labeled DNA ladder, MGA3; *B. methanolicus* MGA3, PB1; *B. methanolicus* PB1, C; control (PCR amplicon of *B. subtilisyrhE* gene).

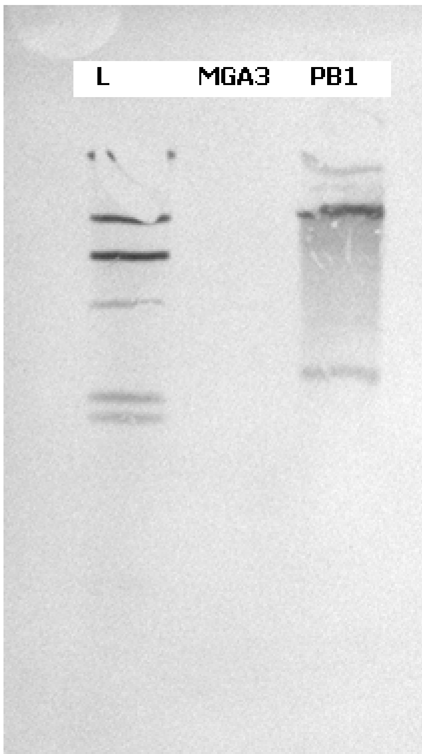


Figure 4.10. Hybridization of *G. kaustophilus* putative *fdh* gene as probe against *B. methanolicus* MGA3 and PB1. L; Dig-labeled DNA ladder, MGA3; *B. methanolicus* MGA3, PB1; *B. methanolicus* PB1.

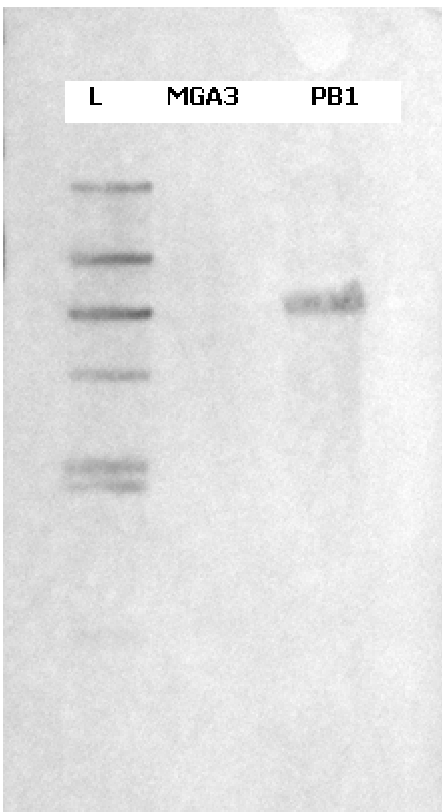


Figure 4.11. Hybridization of *B. subtilis citZ* gene against total DNA from *B. methanolicus* strains MGA3 and PB1. L; Dig-labeled DNA ladder, MGA3; total DNA from strain MGA3, PB1; total DNA from strain PB1.

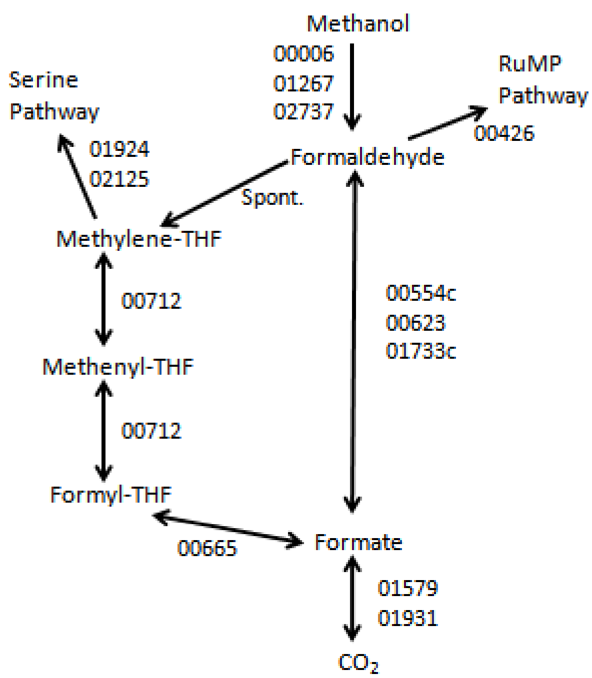


Figure 4.12. Proposed routes for the fate of methanol carbon based on the *B. methanolicus* MGA3 (NCSU) genomic data. ORFs encoding for the putative enzymes are indicated with their respective numbers.

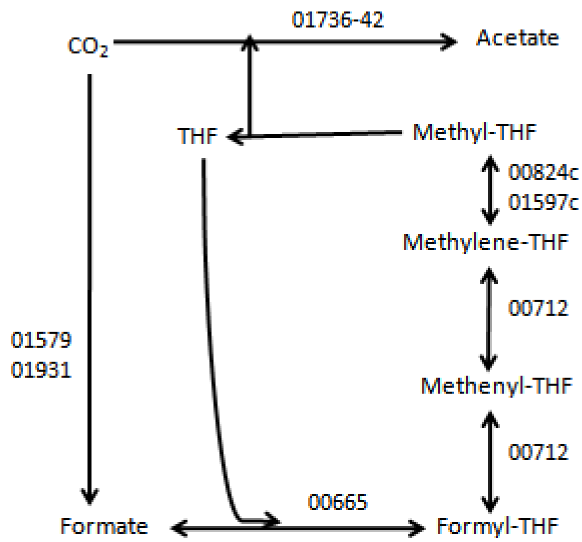


Figure 4.13. Construction of Wood-Ljungdahl pathway from the genomic data. ORFs encoding for the putative enzymes are indicated with their respective numbers.

CHAPTER 5

REGULATION OF METHANOL AND FORMALDEHYDE METABOLISM BY MeOH, FORMALDEHYDE AND FORMATE

Introduction

B. methanolicus is a restrictive methylotroph that is able to grow on a few sugar compounds other than methanol. Arfman *et al*., (1992) summarized the compounds that *B. methanolicus* species can utilize. All *B. methanolicus* species can grow on mannitol at comparable levels with methanol. Although they report data regarding growth on sucrose (Arfman *et al*., 1992), there are no published data regarding growth of *B. methanolicus* on glucose to date. Recently, Heggeset *et al*., suggested (2012) that *B. methanolicus* may also be able to grow on glucose and this was shown in Chapter 2. However, mannitol was the choice of sugar when *B. methanolicus* was intended to grow non-methylotrophically. Mannitol was also the choice of sugar as control when the gene expression of methylotrophic pathways were analyzed and compared to non-methylotrophic gene expression (Jacobsen et. al., 2006).

The regulation of C_1 assimilation in the serine cycle methylotroph *Methylobacterium extorquens* AM1 is achieved via a LysR-type transcriptional regulator that is a homolog to CbbR (Kalyuzhnaya and Lidstrom, 2003). Kalyuzhnaya and Lidstrom (2003) have shown that the *qscR* mutant was not able to grow on C_1 compounds; the activity of the key enzyme of the serine cycle, serine-glyoxylate aminotransferase, was lacking in the qscR mutants. The binding of QscR upstream of most serine cycle enzymes was shown via gel retardation assays in the same study.

There is no study to date to report the regulation of methylotrophy in *B. methanolicus*. The existing data on the differential expression of genes in *B. methanolicus* is based on the growth of the microorganism on methanol and mannitol (Jacobsen *et al*., 2006, Heggeset *et al.*, 2012). Jacobsen *et al*., (2006) reported expression levels of RuMP pathway genes using Real Time-qPCR. They compared the expression levels of *B. methanolicus* MGA3 genes grown on methanol or mannitol.As mentioned, two of these genes, *hps* and *phi*, are known to be on the chromosome while the rest are plasmid-bound. The pBM19 plasmid was shown to be lost after prolonged incubation of *B. methanolicus* MGA3 on mannitol (Brautaset *et. al.*, 2004).

There is a need for a regulated promoter to be used in *B. methanolicus* as a candidate industrial microorganism for the production of amino acids and fine chemicals from methanol. The reported expressed genes, either homologous (Jacobsen et al. 2006 and 2009, Brautaset et al. 2010) or heterologous (Nilasari et al. 2012), used the promoter P_{mdh}. Jacobsen et al. (2006) expressed genes of the carbon assimilation pathway under the P_{mdh} on the shuttle vector pHP13 in order to achieve higher carbon flux through the end product of L-lysine. However, Nilasari *et al.* used another vector pNW33N to successfully transform *B. methanolicus* and express heterogeneous GFP. These studies were constitutive, i.e unregulated. The GFP protein was expressed regardless of the substrate the microorganism was grown on (Smith, 2007).

This chapter investigates the regulation of methanol dissimilation pathway genes that are mentioned in Chapter 4 in response to a spike of methanol, formaldehyde and formate by

RT-qPCR. The strengths of possible promoters to be used in *B. methanolicus* are also assessed.

Materials and Methods

Growth conditions of *B. methanolicus*:*B. methanolicus* MGA3 was grown as described in Chapter 2 in either $MTYM_{MeOH}$ or $MTYM_{MAN.}$The OD_{600} of culture to be spiked was determined experimentally to have sufficient cells (exponentially growing) for RNA extraction. The formaldehyde concentration was also experimentally determined to have sufficient amount to arrest the cell growth for desired time (at least 1 hour) but not to kill the culture. In this regard, combination of OD_{600} of 0.8 and 2 mM formaldehyde was chosen (Figure 5.1). A lower cell density would not be sufficient for desired RNA amount and higher cell density would remove formaldehyde from the media faster. Similarly, a lower formaldehyde concentrarion would be removed from the media faster and higher concentration would kill the cells. The OD_{600} of the culture remained the same during sampling for RNA extraction (Figure 5.2).Thus, exponentially growing cultures were spiked with methanol (100 mM), formaldehyde (2 mM) or formate (2 mM) at $O.D_{600}$ ~0.8. The total RNA was extracted 20 minutes after the spike unless otherwise stated.

Determination of cell lysis and sporulation: DNA concentration in the spent media was measured as an indication of cell lysis for a period of 1 hour after spiking with formaldehyde at above mentioned conditions. Standards of DNA was prepared from Lambda DNA by diluting in MTYM media. A standard curve was constructed using the dilutions of 5ng/ml to 250 ng/ml (Figure 5.3). One milliliter of culture was centrifuged for 1 min at 16,000 g. DNA in the supernatant was measured using NanoDrop-1000. No significant cell lysis was detected during the time frame the RNA extractions were carried out (Figure 5.4)Malachite Green was used to detect sporulation of the cells upon formaldehyde addition. No spore formation was detected for the period of 1 hour past formaldehyde addition.

Extraction of total RNA: Three milliliters of culture was directly transferred into a falcon tube containing 6 ml of Qiagen RNAprotect® Bacteria Reagent (Cat. 76506) and incubated for 5 minutes at room temperature. Then, cells were centrifuged at 10,000g for 5 minutes and the pellet was re-suspended in 0.2 ml TE containing 20 mg/ml lysozyme (Sigma cat. L6876). The sample was then processed according to the Qiagen RNeasy Mini Kit manual. Extracted RNA was quantified using a NanoDrop-1000 and aliquoted into tubes for single use.

Primers and Standards: Primers used for qPCR (listed in Table 1) were designed using software Primer Express 3.0 of Applied Biosystems to be 120-150 bp long and have similar annealing temperatures (~60 °C). For absolute quantification, standards were synthesized by PCR using the Expand High Fidelity PCR System. The reaction mixture contained 400 nmol of primers, 400 µmol of dNTP mix, 0.5 µl of enzyme mix, 1.5 mM $MgCl_2$ and 50-200 ng of total DNA of *B. methanolicus* MGA3. The reaction conditions were: 5 min initial denaturation at 94 °C then, 25 cycles of 1 min annealing at 52 °C, 1 min extension at 72 °C, and 30 sec denaturation at 94 °C. A final extension at 72 °C was added at the end of the cycle. The PCR product was run on an 0.8 % agarose gel and purified by using the Wizard® SV Gel and PCR Clean-Up System (Promega cat. A9281). Purified PCR product was quantified using NanoDrop-1000 and the concentration was calculated based on the molecular mass of the amplicon. A calculated amount of sample was then serially diluted for the standards with the range of 10^{10} to 10^5. Five of these serial dilutions were used to obtain a standard curve depending on the sample to be analyzed.

Table 5.1. List of primers used for RT-qPCR. Underlined sequences are sites for respective restriction sites.

Gene	**Sequence**	**Reference**
Formate THF-ligase(ORF 000 91)	Fw:ATGCTCTGGCAGCTTTAATTGATA Rev: ATGAACAGGTCCGCCTAAACC	This study
Aldehyde dehydrogenase (ORF 00554)	Fw: TGCCTCTACTGCGCGATCTA Rv: GAAGCAGGAAAGCGTGTAGGA	This study
Aldehyde dehydrogenase ORFO 00623)	Fw: AAATGTCCGCCAGTGAAAGAA Rv: TTTGTTTCCCTAATGGGTTTTCC	This study
Methylene tetrahydrofolate dehydrogenase/ cyclohydrolase (ORF 00712)	Fw:GAAGGAGCCGTTGTAATTGATGTA Rv: CGTAATCGTCATCGGTCCAA	This study
Formate dehydrogenase alpha chain (ORF 01579)	Fw: TACGAACGGCTTTGAAGAACTGA Rv: GCATCGGCAAGCAGTTTTG	This study
Methylene tetrahydrofolate reductase (ORF 0 1597)	Fw: TGGCGTTTGGATAAGCTGATAA Rv: TAGAGGGCCTGGGAGCAGAT	This study
CO dehydrogenase large subunit (ORF 01738)	Fw: AGAATTGATACGTCGGCTGCTT Rv: CACCATCTGTTTGTCGCCATT	This study
Formate dehydrogenase (ORF 01931)	Fw: TGTTGCTCCTGAAGCCAATAAAG Rv: AAAGGATGATGTTGCAGGAAGTACA	This study
Pyruvate formate lyase (ORF 02694c)	Fw: CCCTGGTCCATGTGAGGTAATC Rv: TCCGACTGAAGCTACAACAAAATT	This study
Methanol dehydrogenase (ORFs 01267 and 02737)	Fw: TGAAGGTGTCGATGTATCAAAAGAA Rv: TTCACTTTGCGTTCAGTATCTGTGA	This study
Superoxide dismutase (ORF 00650)	Fw: AAACAATGAACATCCATCACACAAA Rv: GCATCAAGGTTGGAAATGACTTCT	This study
mdh (plasmid borne)	Fw: CAAACGTGTGGCATATTAACTGAGT Rv: GGCAGGAGTGGCATTTAACAAC	This study
mdh (chromosomal)	Fw: TGAAGGTGTCGATGTATCAAAAGAA Rv:TTCACTTTGCGTTCAGTATCTGTGA	This study
Pfk	Fw: AAGTGCCATCTCCACCAATC Rv: CCAGGAATGAACGCTGCTAT	Jacobsen et al.
P_{pfk} promoter	Fw: ATGGTACGATCGCAAAGCAAACGG TAC(PvuI) Rv:ATGGTAGAATTCTCTTGTTTTCCTCCATATC(EcoRI)	This study
Alu156	Fw:ATGGTAGAATTCAAGATTATTATGCCACAC(EcoRI) Rv:ATGGTACGATCGCCCGGATCCGTCCTTCC (PvuI)	This study
repB	Fw: GTCTGTTCGTCCTTGGACTT Rv: AATGCTTGGAGCCGAGGATA	Jacobsen et.al.,
Gfpuv (qPCR)	Fw2: TGTTCAATGCTTTTCCCGTTATC Rv2: TTGACTTCAGCACGCGTCTT	This study

Table 5.1 Continued

GFPuv (Cloning)	Fw:ATGGTAGAATTCATGAGTAAAGGAGAAG(EcoRI) Rv:ATGGTAAAGCTTTTATTTGTAGAGCTCATC(HindIII)	This study
hps	Fw: CCTTGTTGACATGATCGCAGTT Rv: AATGGGTTTTTACCTACTGCTTGAA	This study
G. kaustophilus 2797314-2798810	Fw:ATGGTAGAATTCATGACAATGATTTCGGCG(EcoRI) Rv:ATGGTAGGATCCTTAAAACAACGTCATCGG(BamHI)	This study
P_{mdh} promoter	Fw:GTAGAATTCTCTTACTACCTCCTATTTATG(EcoRI) Rv: ATGGTACGATCGAATCGTTCATTAAAGAGC(PvuI)	This study
Hps+phi	Fw:ATGGTAGAATTCATGGAACTTCAATTAG(EcoRI) Rv:ATGGTAAAGCTTCTACTCAAGATTAGC(HindIII)	This study

RT-qPCR: The reactions were conducted with an Applied Biosystems 7300 Real-Time PCR system and a Quantitect SYBR Green PCR Kit was used for qPCR experiments. For absolute quantification, both reverse transcription and qPCR were conducted in the same tube. The plate contained the samples and the standards with 5-log range for the standard curve. The reaction mixture contained 25 µl of PCR mix with SYBR Green, 200 nmol of forward and reverse primers, 50 ng of total RNA as template, and 0.5 µl of Reverse Transcriptase (RT) mix in a total of 50 µl. The RT mix was omitted from the standards. The reaction conditions were as follows; 30 minutes at 50 °C for RT reaction, 15 min at 95 °C for activation of Taq polymerase and then 30 cycles of 30 sec. denaturation at 95 °C, 30 sec. annealing at 55 °C and 30 sec. extension at 72 °C and then dissociation curve from 60 °C to 95 °C with 1 °C increment at every 15 sec. while recording fluorescence. The results were normalized using 16S rRNA.

One-step RT-qPCR, the reverse transcription and qPCR reactions in the same tube, was used for relative quantification. The conditions of the one-step RT-qPCR are the same as the absolute quantification. Relative expression of genes in methanol grown cells was assessed with respect to mannitol grown cells as the calibrator. Both *repB* and 16S rRNA genes were used as endogenous controls mainly to compare the results with previous reports. The results are presented as relative to counterparts of genes from mannitol grown cells.

Construction of GFPvectors: The *E.coli*-*B.subtilis* pHP13 shuttle vector was used as the backbone for cloning genes into *B. methanolicus* MGA3. The propagation and transformation of the vector was conducted in *E. coli* as described by Maniatis et. al.(1982). The transformants were screened via PCR for positive colonies. Upstream (500 bp) sequences of *mdh,* and *pfk* were amplified by PCR using the primers shown in Table 5.1 and cloned into pHP13 via respective restriction sites. The Alu156 promoter was synthesized by Integrated DNA Technologies (Coralville, IA, US) as single strand and then was amplified using the primers Alu156Fw and Alu156Rv. The GFPuv gene was amplified using PCR and cloned downstream of promoters of P_{mdh},$P_{pfk,}$ and $P_{Alu156.}$ The sequences of the clones were confirmed by sequencing the vectors with custom designed and commercially available primers. The pNW33N (Nilasari et al., 2012) vector was kindly provided by Dr. Claire Komives and it is labeled as P_{mdhK} for distinction from the P_{mdh}. It contains 11 amino acids of methanol dehydrogenase before the GFPuv.

Construction of pTS123 and pHPS12: The P_{mdh} promoter was amplified by PCR and cloned into pHP13 using respective restriction enzyme sites. The putative *faldh* of *G. kaustophilus* (ORF 2797314-2798810) was also PCR amplified using primers shown in Table 5.1 and cloned downstream of P_{mdh}. This vector is called pTS123. Similarly, the P_{pfk} was also cloned into pHP13. Genes of *hps*and *phi* were amplified by PCR monocistronically using the primers in Table 5.1 and cloned into downstream of P_{pfk}. This vector is called pHPS12.

Electroporation of *B. methanolicus*: The electroporation of *B. methanolicus* was conducted with the following modifications to published protocols (Jacobsen et al. 2006, Nilasari et al. 2012). The cells were grown in 5 ml of SOBsuc after electroporation for 2 hours in a 50 ml falcon tube and then 10 ml SOBsuc with 10 μg/ml final concentration of choloramphenical was added to the tube and incubated 6 more hours before plating. A colony from the plate was inoculated into 5 ml of MTYM with 60 mM methanol and incubated for 6 hours. The entire volume was then transferred into 50 ml MTYM media in a 250 ml shake flask with 150 mM methanol and 5 μg/ml chloramphenicol. The culture was incubated overnight and frozen stocks were prepared by flash freeze in liquid nitrogen.

Results

Methanol dissimilation pathway genes are down regulated by methanol, formaldehyde and formate: The proposed methanol oxidation pathway derived from the genome sequence of *B. methanolicus* MGA3 is presented in Figure 4.12. To summarize briefly, methanol is oxidized to formaldehyde by methanol dehydrogenases (ORFs 00006, 01267, 02737)in the first step. Formaldehyde is then oxidized to formate either by putative formaldehyde dehydrogenases (ORFs 00554c, 00623c, 01733c) or THF-linked methylene tetrahydrofolate dehydrogenase/cyclohydrolase (ORF 00712) and formyl-THF ligase (ORFs 00665).

Regulation of methanol oxidation: Figures 5.5-5.9 depict the expression of *mdh* genes in response to methanol addition. As can be seen from the figure 5.5, the expression level of pBM19 *mdh* stays stable over the course of 60 minutes after methanol addition when *B. methanolicus* MGA3 was grown in MTYM methanol media. However, the expression is significantly down-regulated when mannitol grown *B. methanolicus* MGA3 was exposed to a 100 mM methanol spike, as can be seen in Figure 5.6. The expression level of *mdh* was decreased almost one-log in 20 minutes. The expression of chromosomal *mdh* genes in response to methanol addition is different from pBM19 *mdh*in methanol-grown *B. methanolicus* MGA3 in that it was down regulated immediately (until 10 minutes post-spiking) and then starts to recover back although, the expression level after 60 minutes is still significantly lower than pre-spiking level. The expression of pBM19 *mdh* is also down-regulated when methanol or mannitol grown *B. methanolicus* MGA3 was subjected to formaldehyde (2 mM) and formate spike (2 mM). As seen in Figures 5.9 and 5.10, pBM19 *mdh* was down-regulated in response to a formaldehyde and formate spike although formaldehyde was more potent to down-regulate the gene.

An important observation from the expression levels of *mdh* is that it is expressed constitutively both in mannitol and methanol-grown *B. methanolicus* MGA3. This result is consistent with the expression of GFP under P_{mdh} in both methanol and mannitol grown cells (Nilasari et. al., 2012). The expression levels of pBM19 *mdh* and chromosomal *mdh* genes are comparable between methanol and mannitol grown cells. However, there are clear differences in expression levels of pBM19 and chromosomal *mdh* genes. The expression level of pBM19 *mdh* is one-log higher than the combined expression of *mdh* genes from ORFs 01267, 02737 (the qPCR primers for chromosomal *mdh* genes amplify both genes). This might be because of the copy number of pBM19, ~10-16 copies per cell (Brautaset et. al., 2004).

Regulation of formaldehyde oxidation: The second step in methanol dissimilation is the conversion of formaldehyde to formate which can be either through THF-linked enzymes or formaldehyde dehydrogenases. The expression levels of putative genes for these pathways are assessed via RT-qPCR.

Figures 5.11 and 5.12 depict the regulation of putative aldehyde dehydrogenase (possibly formaldehyde) in response to formaldehyde, formate and methanol spiking. Both putative aldehyde dehydrogenases are down regulated in response to spikes by methanol and its metabolites. However the down-regulation is more profound in methanol-grown cells.

Figures 5.13 and 5.14 summarize the regulation of putative THF-linked enzymes as a second possible pathway for the oxidation of formaldehyde to formate. The putative methylene tetrahydrofolate dehydrogenase/cyclohydrolase (ORF 00712) is down regulated by methanol, formaldehyde and formate in both methanol and mannitol grown cells. The putative formyl-THF ligase enzyme (ORF 00665) is also down regulated by these compounds. However, the effect of formate is more profound on Formyl-THF ligase than formaldehyde while it is reversed for methylene tetrahydrofolate dehydrogenase/cyclohydrolase.

Regulation of formate dehydrogenase: The third step in the dissimilation of methanol is the oxidation of formate to CO_2. This reaction is catalyzed by formate dehydrogenase. Two putative formate dehydrogenases are ORFs 01579 and 01931. Figures 5.15 and 5.16 summarize the regulation of these genes in response to methanol and its metabolites. Both of the formate dehydrogenases are down-regulated in response to spiking of methanol and its metabolites. Although expression levels of ORF 01931 are comparable between methanol and mannitol grown cells, the ORF 01579 in mannitol-grown cells is expressed as twice the level that it is expressed in methanol-grown cells.

Regulation of methanol assimilation (RuMP) pathway gene *hps*: Figures 5.17 and 5.18 summarize the regulation of *hps* in response to methanol and its metabolites. The *hps* gene encodes for the first enzyme of the RuMP pathway that draws the formaldehyde to assimilative pathways. It is surprising that the expression level of *hps*(7.42×10^{-1}) which is located in the chromosome, is significantly higher than the expression levels of pBM19 *mdh* (Met-5 sample 2.23×10^{-1}, see Figures 5.17 and 5.5). The *hps* gene in methanol grown cells is up-regulated approximately 3-times compared to mannitol grown cells (7.42×10^{-1} vs 2.15×10^{-1}). Although not at the same level, this result confirms the upregulation of *hps* reported by Jacobsen et. al., (2006). However, the *hps* is also down regulated when methanol-grown cells are subjected to formaldehyde and formate spikes. Despite this down-regulation in methanol-grown cells, it is clearly upregulated when mannitol-grown cells are subjected to methanol, formaldehyde or formate. Almost 2-fold (2.15×10^{-1} vs 4.17×10^{-1}**)** up-regulation is evident in Figure 5.18.

Regulation of pfk: Another RuMP pathway gene investigated was the phosphofructokinase (*pfk*), which is located on the pBM19 plasmid. It catalyzes the next reaction after *hps* and *phi* in the RuMP pathway; conversion of fructose-6 phosphate to fructose-1,6-biphosphate. Figure 5.19 depicts its regulation in response to formaldehyde. The upregulation of *pfk* in methanol grown cells is apparent in Figure 5.19 which confirms the upregulation reported before (Jacobsen et. al., 2006) although at a significantly lower level. Despite being encoded on the pBM19 plasmid, the expression level of *pfk* is significantly lower than *hps* when figures 5.17 and 5.19 are compared. The *pfk* gene is upregulated in methanol-grown cells compared to mannitol grown cells and spiking mannitol-grown cells with formaldehyde caused about 2-fold upregulation. Although small, spiking methanol-grown cells with formaldehyde also caused an upregulation of *pfk*.

Regulation of Wood-Ljungdahl pathway enzymes: The Wood-Ljungdahl pathway was described in Chapter 4. The THF-linked formaldehyde oxidation enzymes methylene tetrahydrofolate dehydrogenase/cyclohydrolase (ORF 00712) and formyl-tetrahydrofolate ligase (ORF 00665) along with formate dehydrogenase (ORFs 01579, 01931) also function in Wood-Ljungdahl pathway as depicted in Figure 4.13. The regulation of these genes in response to methanol and its metabolites was discussed above. The expression of CO dehydrogenase large subunit (ORF 01738) and methylene tetrahydrofolate reductase (ORF 01597) are also investigated. Figure 5.20 shows that the expression levels of putative methylene tetrahydrofolate reductase in methanol and mannitol-grown *B. methanolicus* are comparable and is down regulated in both samples. However, the expression level of CO dehydrogenase (ORF 01738) in mannitol-grown cells is higher than in methanol grown cells while the gene is also down-regulated by methanol and its metabolites (Figure 5.21).

Superoxide dismutase: The methanol metabolite formaldehyde is a reactive oxygen species, and it reacts nonspecifically with bio-molecules. Accumulation of formaldehyde alone may cause an oxidative stress which can result in down regulation of metabolism as a whole except for the stress response genes. Therefore, the expression level of putative superoxide dismutase (ORF 00650) was investigated. Figure 5.22 shows the response of putative superoxide dismutase in response to formaldehyde and formate spikes. As seen in the figure, the expression level of superoxide dismutase is comparable between the methanol-grown and mannitol grown cells. A spike of formaldehyde did not cause any down or upregulation of superoxide dismutase in methanol-grown cells, although there is a slight upregulation in mannitol-grown cells. However, a spike of formate caused down regulation only in methanol-grown cells, not in the mannitol grown cells. This result suggests that the down regulation of genes discussed above was not due to a global oxidative stress in response to formaldehyde.

Evaluation of promoter strength for regulated protein expression in *B. methanolicus*: Regulation of recombinant protein expression is an important aspect of recombinant protein production and pathway engineering. In this regard, well established systems are available especially for *E. coli* like pET expression systems from Clontech labs (Mountain View, CA) which utilizes bacteriophage T7 RNAP and the T7 promoter for induction of transcription. The T7 expression system was also engineered for expression in *B. subtilis* (Chen et. al., 2010). A methanol induced promoter is desirable for a possible *B. methanolicus* expression system. The P_{mdh} of *B. methanolicus* has been used for both homologous and heterologous gene expression as discussed above. To elucidate the behavior of P_{mdh} and compare with $P_{pfk}$and P_{Alu156}, the transcription level of the reporter GFPuv was assessed. Figure 5.23 summarizes the expression of GFPuv by these promoters. As seen in the figure, the expression level of GFP under P_{mdhK} is comparable to the expression level of pBM19 *mdh*, and it is also down regulated by a spike of formaldehyde. The expression level underthe control of P_{mdh}, P_{pfk} and P_{Alu156} are significantly lower than the expression level under P_{mdhK}. However, there is a slight upregulation in P_{pfk} regulated expression with a formaldehyde spike which is again consistent with upregulation of *pfk* with formaldehyde spike in mannitol grown cells. The difference in expression levels of P_{mdhK} and P_{mdh} is most probably due to introduction of 6-bp between the promoter region and start codon during vector construction.

Formaldehyde tolerance of *B. methanolicus* MGA3 carrying pTS123 and pHPS12:The vectors pTS123 and pHPS12 carry genes for enzymes responsible for removal of formaldehyde via oxidation to formate and assimilation into RuMP pathway, respectively. Growth inhibition experiments were conducted to test the formaldehyde tolerance of recombinant *B. methanolicus* MGA3, as described in Chapter 2. Figure 5.25 demonstrates the formaldehyde tolerance of *B. methanolicus* MGA3 carrying putative *G. kaustophilus faldh* under P_{mdh}. As can be seen from the figure, recombinant *B. methanolicus* MGA3 is able to tolerate formaldehyde spikes up to 1 mM. Similarly, recombinant *B. methanolicus* MGA3 carrying extra copies of *hps* and *phi*genes under P_{pfk} is also able to tolerate 1 mM formaldehyde as can be seen on figure 5.26 (See figure 2.6 for comparison).

Conclusions

B. methanolicus is a restrictive methylotrophic microorganism that is also capable of utilizing mannitol. As a candidate industrial microorganism, understanding the regulation of methanol metabolism, especially methanol dissimilation, is crucial for possible manipulations since the substrate carbon is lost to the atmosphere as CO_2. There exist reports comparing the expression profiles of methanol and mannitol-grown cells (Jacobsen et. al., 2006, Heggeset et.al., 2012), but these do not answer the question of how methanol and formaldehyde metabolism are affected by methanol, formaldehyde and formate.

I employed RT-qPCR to elucidate the regulation of methanol and formaldehyde metabolism, especially the methanol dissimilation routes, in response to methanol, formaldehyde and formate, but also included some of the methanol assimilation genes for reasons of comparison with the published reports which gave important insights into the regulation of the metabolism. First of all, my results agree with the report of Jacobsen et. al, (2006) in that the *hps* and *pfk* genes are expressed at higher levels in methanol-grown cells. However, expression of pBM19 *mdh* does not agree with their report that they are comparable in methanol and mannitol-grown cells. This is most probably due to the difference of quantification; absolute vs relative. They used pBM19-borne *repB* to normalize the expression levels which may cause miscalculation since it is known that pBM19 is lost from *B. methanolicus* MGA3 after prolonged incubation on mannitol (Brautaset et. al., 2004). However, in the experimental settings used in this study, the *repB* expression, thus copy number of pBM19, is comparable between methanol and mannitol grown *B. methanolicus* MGA3 (Figure 5.24).Considering the regeneration time of *B. methanolicus* MGA3 (1 hour), it is safe to assume that copy number of pBM19 will stay the same for 1-hour period. Additionally, the absolute expression level of *repB* changes when the methanol or mannitol grown *B. methanolicus* MGA3 is spiked with methanol (Figure 5.24).

The presence of two more *mdh* genes on the chromosome was described in Chapter 3. The qPCR primers for chromosomal *mdh* genes were designed to amplify both of these genes. Thus, Figures 5.7 and 5.8 depict combined expression of chromosomal *mdh* genes. The expression of pBM19 *mdh* is one-log higher than the combined expression of ORFs 01267 and 02737. From this we can infer that pBM19 *mdh* is responsible for most if not all of the methanol dehydrogenase activity. Another striking result is the down regulation of the *mdh* genes in response to methanol spike when *B. methanolicus* MGA3 was grown on mannitol. Contrary to expectations, spiking with methanol did cause the down-regulation of *mdh* genes.

The same pattern of down regulation is also true for the putative methanol dissimilation pathway genes. The putative genes for formaldehyde and formate oxidation are down regulated in response to not only methanol but also formaldehyde and formate spikes. Down regulation with formate spike is also interesting since formate does not have any inhibitory effect on the growth of *B. methanolicus* at a concentration of 25 mM (Pluschkell, 1998) as discussed in Chapter 2.

Contrary to the dissimilation pathway genes, the formaldehyde assimilation pathway genes *hps* and *pfk* are upregulated in response to methanol, formaldehyde and formate (Figures 5.17 and 5.18) in mannitol grown cells. The upregulation by methanol and its metabolites are clearly seen in Figures 5.17 and 5.18 after spiking. But, in methanol-grown cells, the *hps* is down regulated from its already high levels. The expression of *hps* in methanol-grown *B. methanolicus* MGA3 is almost twice that compared to mannitol-grown cells. As mentioned above, the *mdh* is expressed constitutively which will result in conversion of methanol into formaldehyde and the formaldehyde assimilation pathway is saturated, while in mannitol grown cells it is idle. Spiking of formaldehyde or formate will cause the already saturated assimilation pathway to overflow in methanol-grown cells which will cause accumulation and toxicity. However, any additional formaldehyde or formate will not cause any accumulation in mannitol-grown cells as the cells can tolerate it. This also explains the interesting result in Chapter 2, Figures 2.6 and 2.7 where mannitol-grown *B. methanolicus* MGA3 is more tolerant to formaldehyde spike than methanol-grown cells.

In the search for suitable promoters for protein expression and pathway engineering in *B. methanolicus*, I assessed the expression level of GFPuv from different promoters. The low expression levels from P_{mdh}, P_{pfk} and P_{Alu156} are most probably due to the 6-bp introduced between the start codon and the promoter regions since the promoter sequence of P_{mdhK} and P_{mdh} are otherwise identical. From this low level of expression it will be misleading to conclude if P_{Pfk} and P_{Alu156} are better than P_{mdhK}. However, as seen from Figure 5.23, P_{mdhK} is capable of transcribing high levels of protein comparable to its native downstream *mdh* although the expression from P_{mdhK} is also down-regulated by formaldehyde. From the expression of *hps*, it would be proper to recommend usage of P_{hps} since it has high expression levels comparable to plasmid encoded genes, and is upregulated in response to formaldehyde.

The overexpression of genes responsible for formaldehyde removal (either by dissimilatory or assimilatory pathways) confers *B. methanolicus* an ability to tolerate higher formaldehyde concentrations. As can be seen from the Figure 5.25, overexpression of putative *G. kaustophilusfaldh* results in higher formaldehyde tolerance when compared to wild type *B. methanolicus* MGA3. Moreover, the growth rate of *B. methanolicus* MGA3 (pTS123) in 150 mM methanol is also comparable to the growth rate of wild type *B. methanolicus* MGA3, 0.547±0.006 h^{-1} and 0.549±0.007 h^{-1}, respectively. However, growth rate of *B. methanolicus* MGA3 (pHPS12) is higher than growth rate of wild type, 0.599±0.004 $h^{-1}$and 0.549±0.007 h^{-1}, respectively, in addition to the higher tolerance of formaldehyde (Figure 5.26). This suggests that more formaldehyde is being assimilated by *B.*

methanolicus MGA3 (pHPS12). Dissimilation of carbon by both recombinant and wild type strains need to be compared to have better picture of formaldehyde assimilation by *B. methanolicus* MGA3.

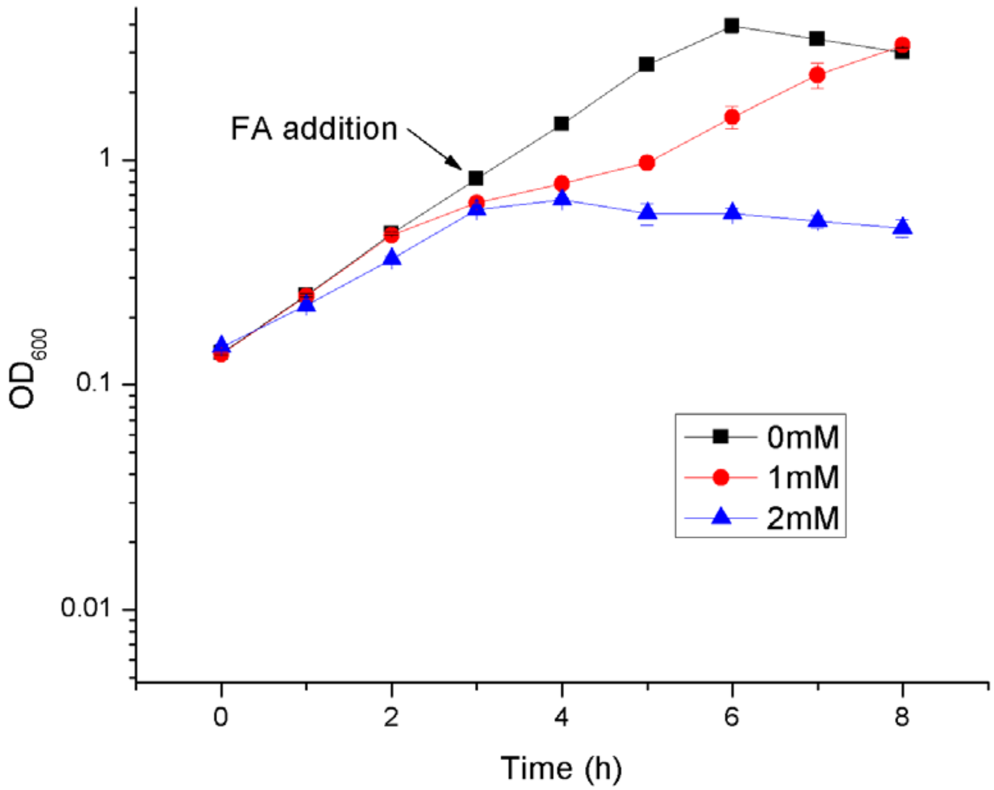


Figure 5.1. Response of $MTYM_{MeOH}$ grown *B. methanolicus* MGA3 to formaldehyde spikes. *B. methanolicus* MGA3 was grown in a 50 ml $MTYM_{MeOH}$ in 250 ml baffled flask and challenged with different concentrations of formaldehyde when the cells were growing exponentially (OD_{600} ~0.8). (n=3)

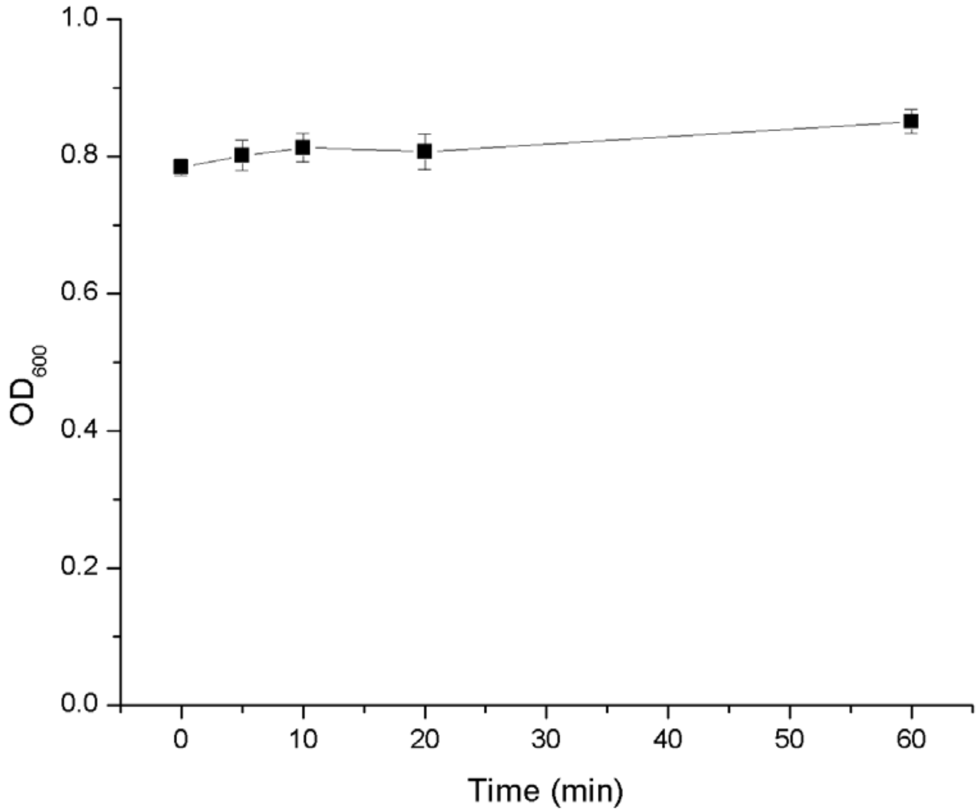


Figure 5.2. Change in OD_{600} of *B. methanolicus* MGA3 after spiking with 2 mM formaldehyde. *B. methanolicus* MGA3 was grown in a 50 ml $MTYM_{MeOH}$ in 250 ml baffled flask and challenged with 2 mM of formaldehyde when the cells were growing exponentially (OD_{600} ~0.8). (n=3)

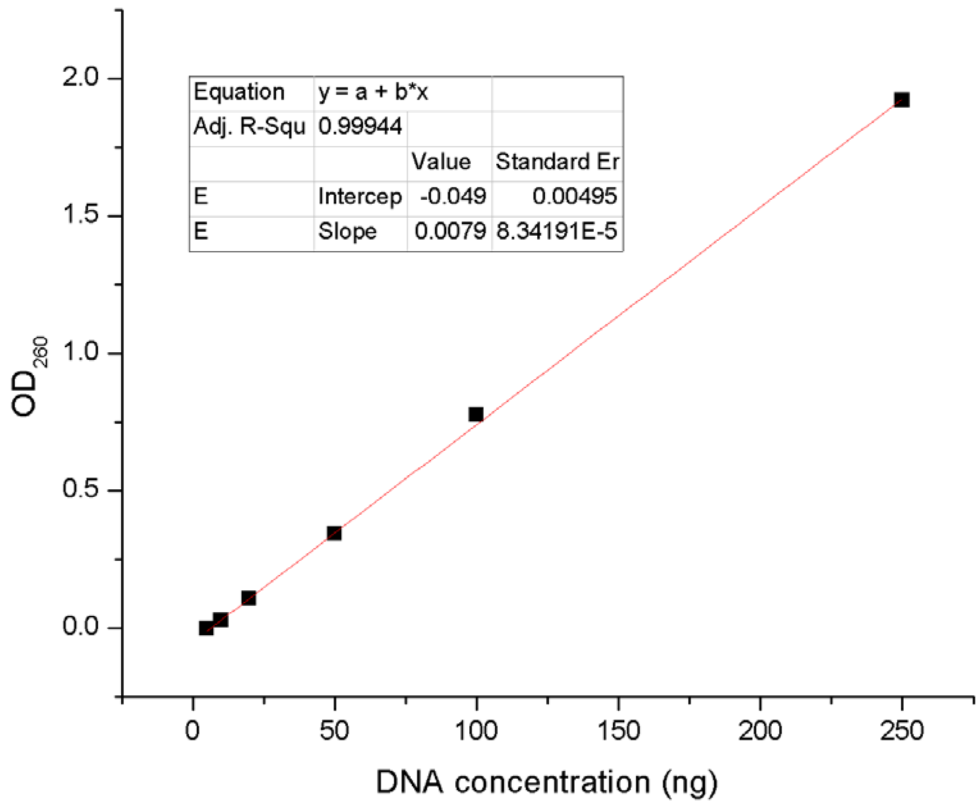


Figure 5.3. Standard curve for calculation of DNA concentration in the culture supernatant.Standards were prepard using Lambda DNA by diluting in MTYM to have a range of 5 ng/ml to 250 ng/ml. The DNA concentration was measured using NanoDrop-1000.

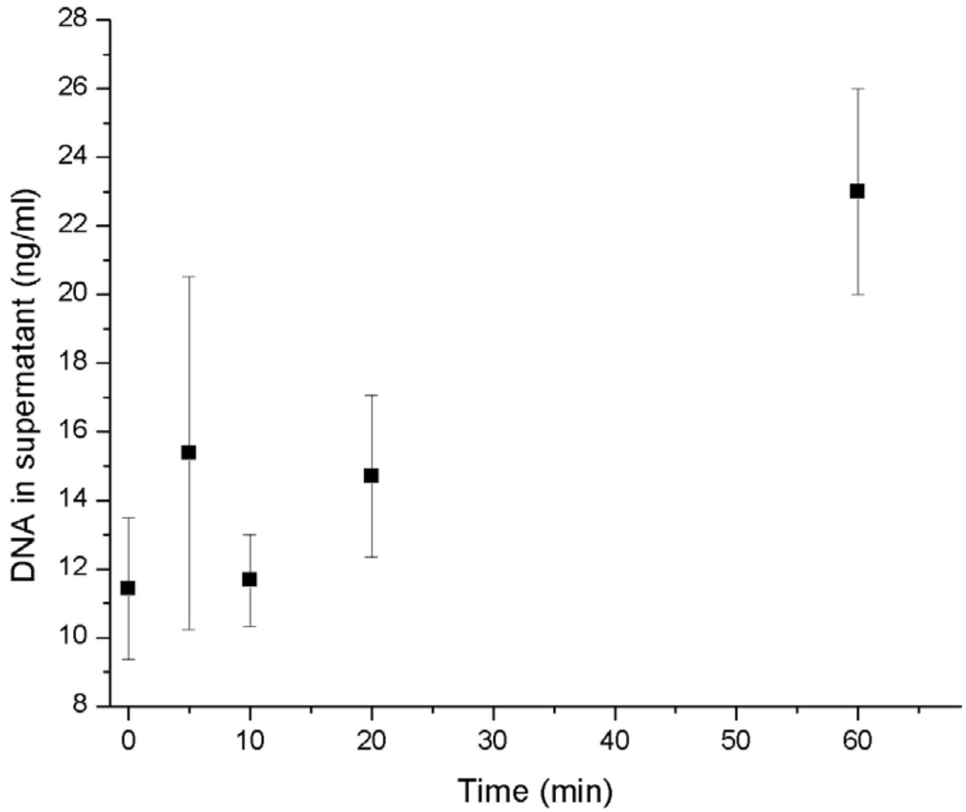
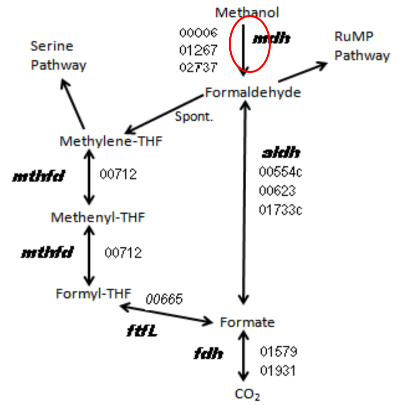


Figure 5.4. DNA concentration in the supernatant of *B. methanolicus* MGA3 after spiking with formaldehyde. The DNA in the supernatant was measured using NanoDrop-1000.



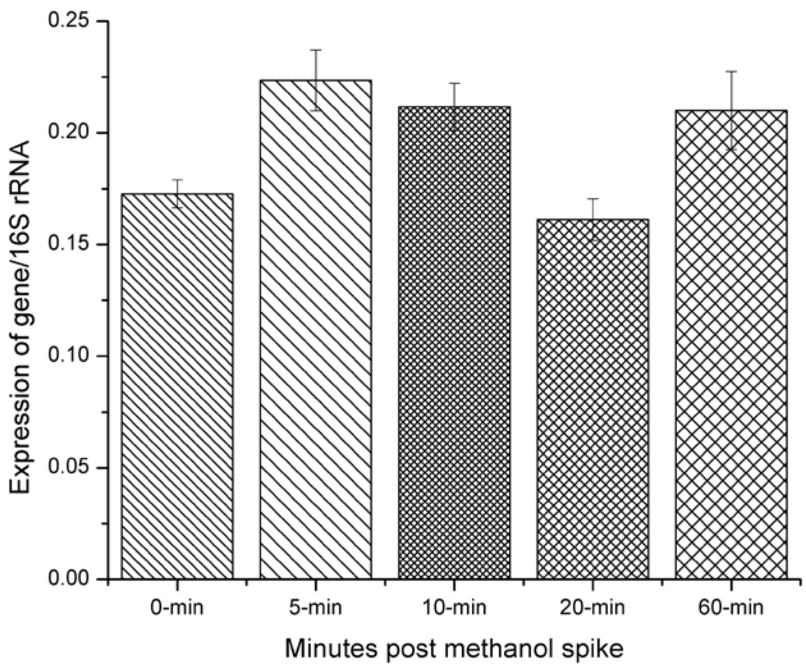
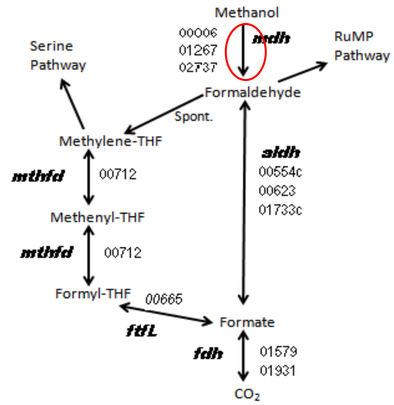


Figure 5.5. Regulation of pBM19 *mdh* in response to 100 mM methanol addition to methanol grown *B. methanolicus*.*B. methanolicus* MGA3 was grown in $MTYM_{MeOH}$ and spiked with methanol at $O.D_{600}$ of ~0.8. Then, the total RNA was extracted as described after 5, 10, 20, and 60 min. The sample “0-min” was taken immediately before the spike. (n=3)



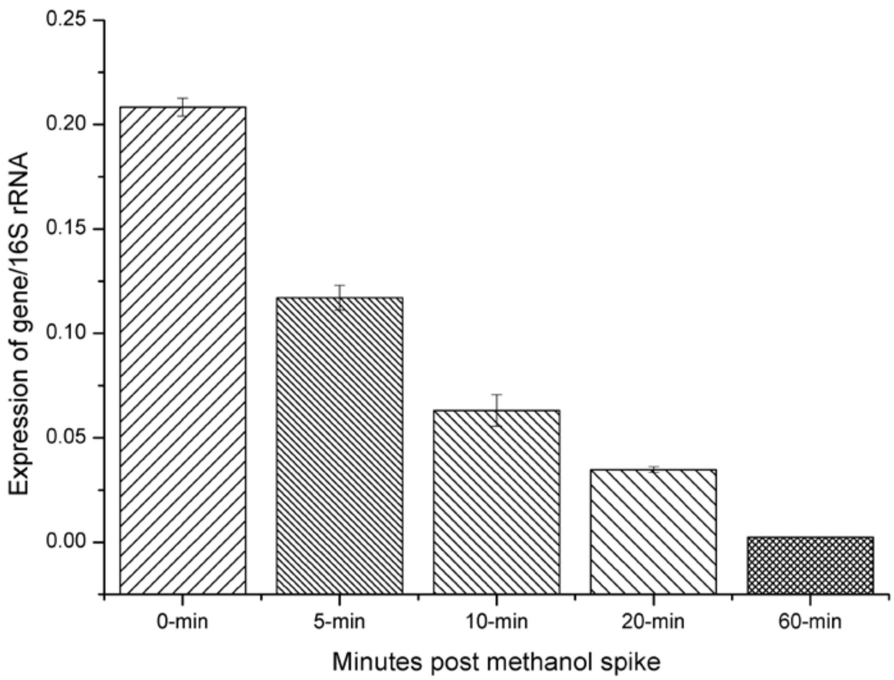
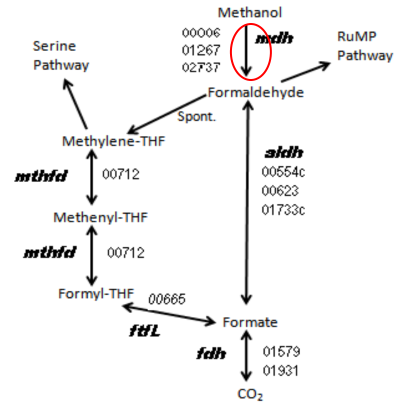


Figure 5.6. Regulation of pBM19 *mdh* in response to 100 mM methanol addition in mannitol grown *B. methanolicus*.*B. methanolicus* MGA3 was grown in $MTYM_{MAN}$ and spiked with methanol at $O.D_{600}$ of ~0.8. Then, the total RNA was extracted as described after 5, 10, 20 and 60 min post spike. The sample "0-min" was taken immediately before the spike. (n=3)



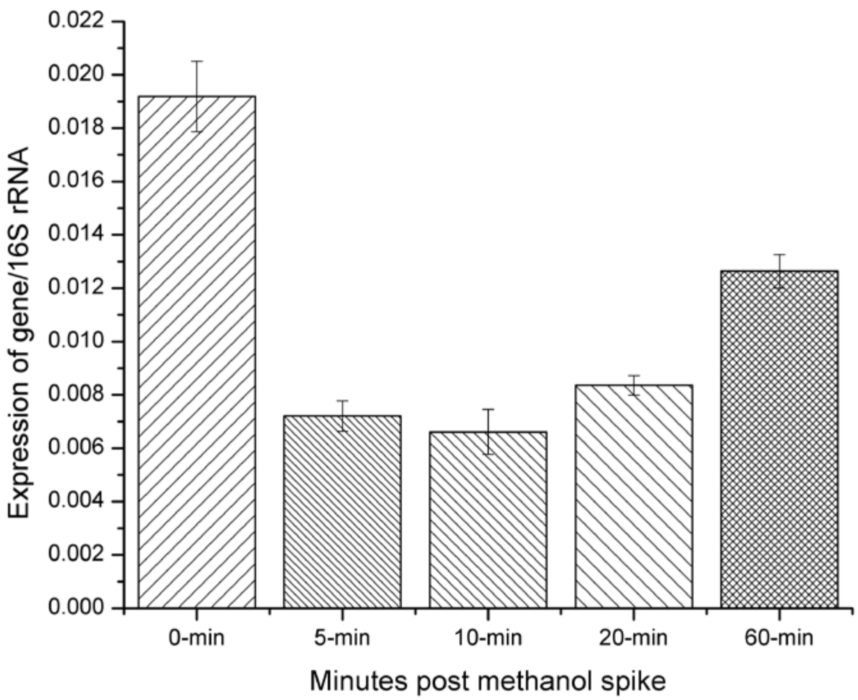
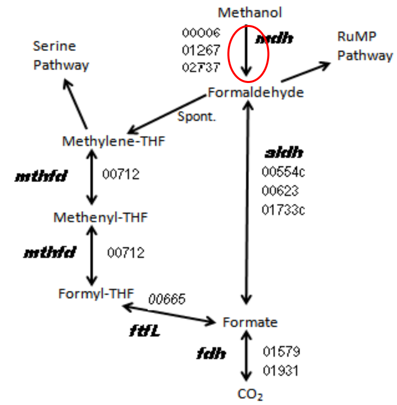


Figure. 5.7. Regulation of chromosomal *mdh* (ORFs 01267 and 02737) genes in response to 100 mM methanol addition. *B. methanolicus* MGA3 was grown in $MTYM_{150}$ and spiked with methanol at $O.D_{600}$ of ~0.8. Then, the total RNA was extracted as described after 5, 10, 20, and 60 min. The sample "0-min" was taken immediately before the spike. (n=3)



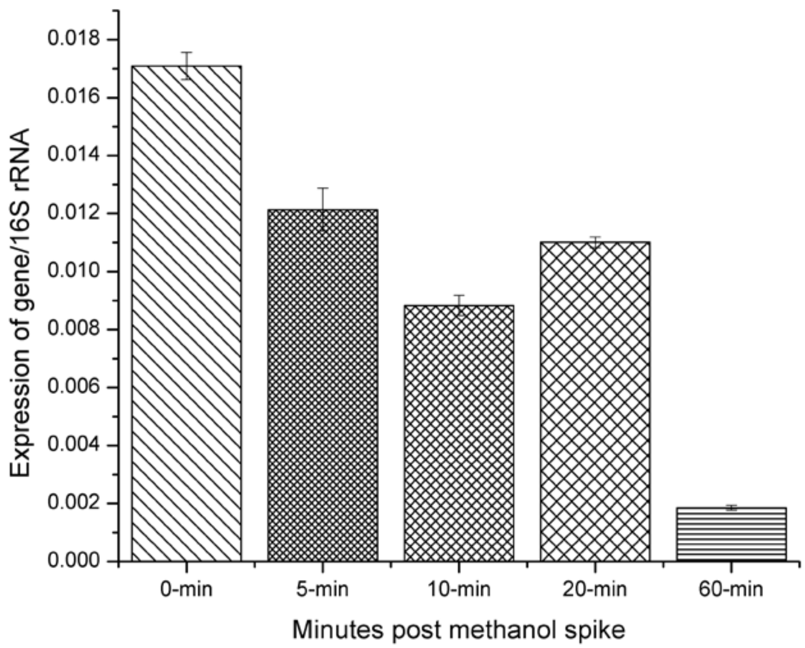
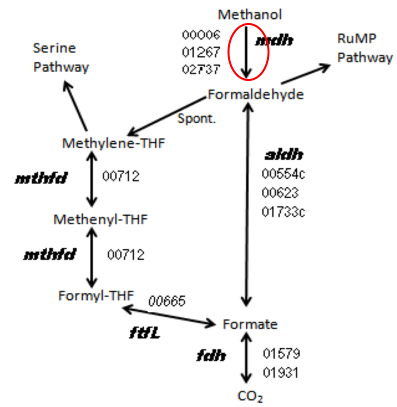


Figure 5.8. Regulation of chromosomal *mdh* genes (ORFs 01267 and 02737) in response to 100 mM methanol addition.*B. methanolicus* MGA3 was grown in $MTYM_{MAN}$ and spiked with methanol at $O.D_{600}$ of ~0.8. Then, the total RNA was extracted as described after 5, 10, 20, and 60 min. The sample "0-min" was taken immediately before the spike. (n=3)



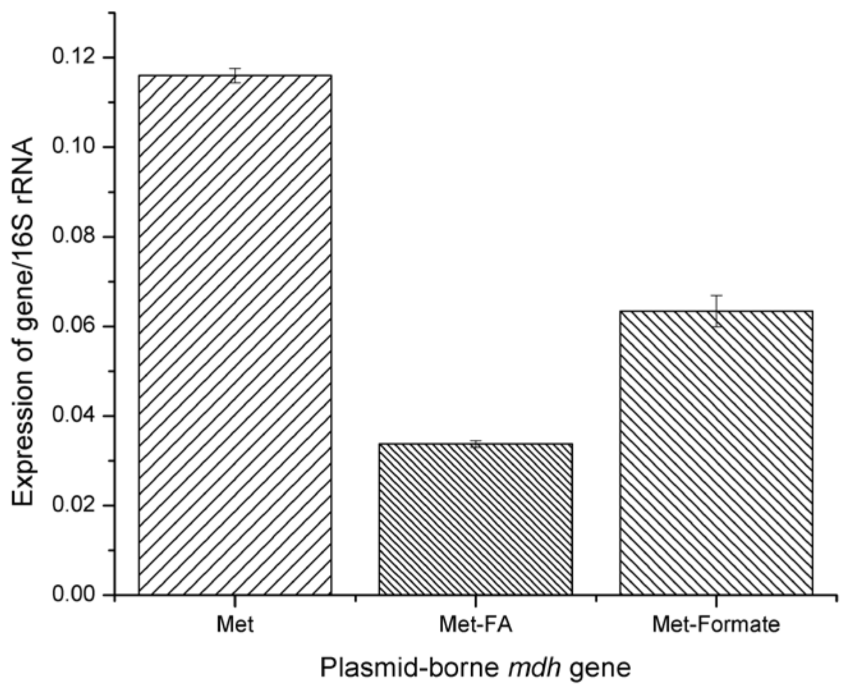
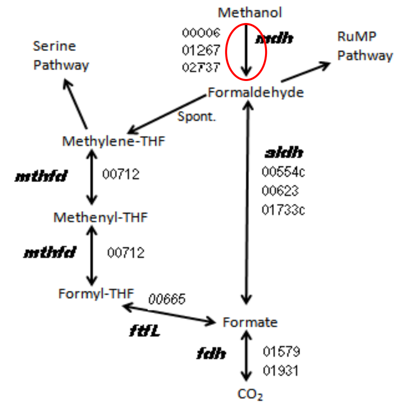


Figure 5.9. Regulation of pBM19 *mdh* in response to formaldehyde (2 mM) and formate (2 mM) addition.*B. methanolicus* MGA3 was grown in $MTYM_{150}$ and spiked when the $O.D_{600}$ is ~0.8. The total RNA was extracted 20 minutes post-spiking. Met; immediately before spike, Met-FA; spiked with formaldehyde, Met-Formate; spiked with formate. (n=3)



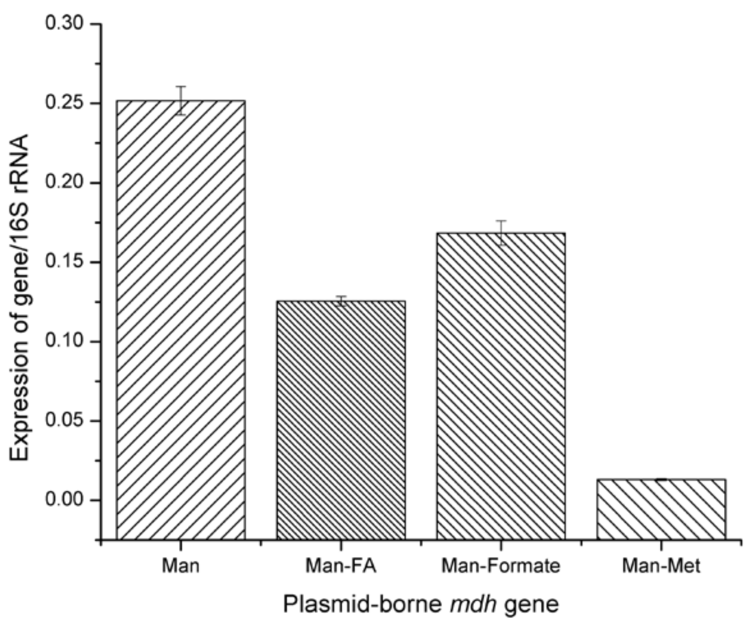
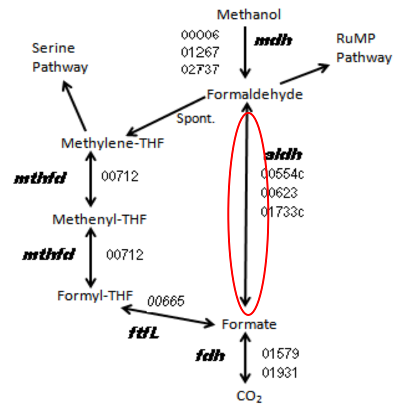


Figure 5.10. Regulation of pBM19 *mdh* in response to formaldehyde (2 mM), formate (2 mM) and methanol (100 mM) addition. *B. methanolicus* MGA3 was grown in $MTYM_{MAN}$ and spiked when the $O.D_{600}$ is ~0.8. The total RNA was extracted 20 minutes post-spiking. Man; immediately before spike, Man-FA; spiked with formaldehyde, Man-Formate; spiked with formate, Man-Met: spiked with methanol. (n=3)



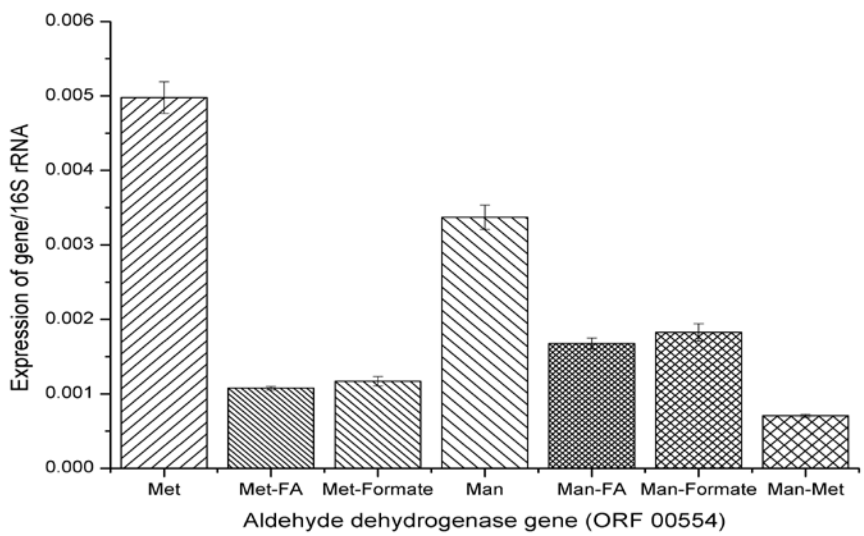
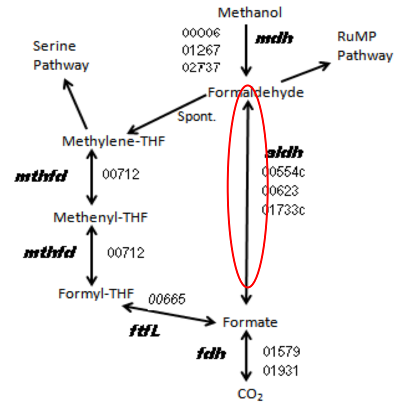


Figure 5.11. Regulation of putative aldehyde dehydrogenase (ORF 00554c) in response to formaldehyde (2 mM), formate (2 mM) and methanol (for mannitol-grown cells, 100 mM). *B. methanolicus* MGA3 was grown in either methanol (Met) or mannitol (Man) and spiked when the $O.D_{600}$ is ~0.8. The total RNA was extracted 20 minutes post-spiking.Met; immediately before spike, Met-FA; spiked with formaldehyde, Met-Formate; spiked with formate, Man; immediately before spike, Man-FA; spiked with formaldehyde, Man-Formate; spiked with formate, Man-Met: spiked with methanol.(n=3)



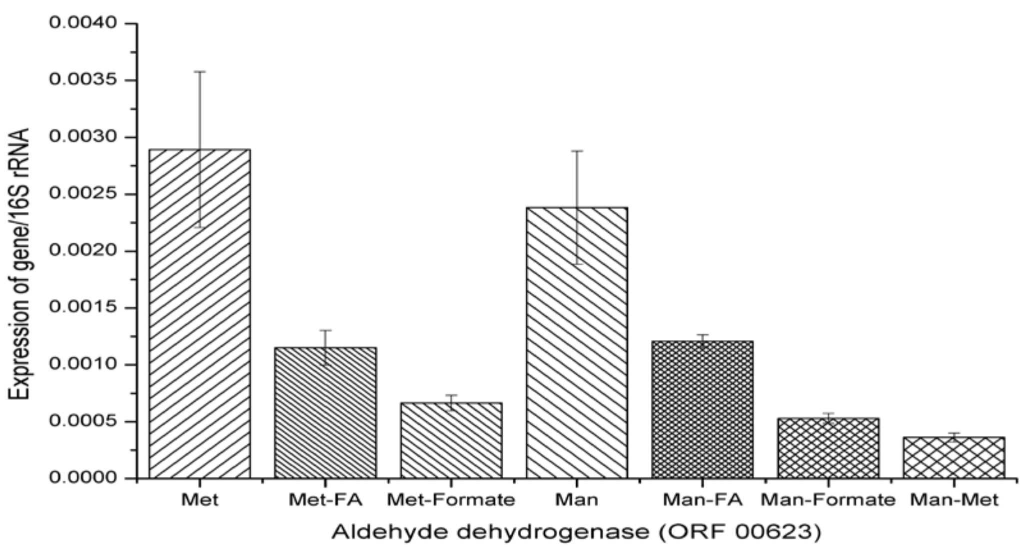
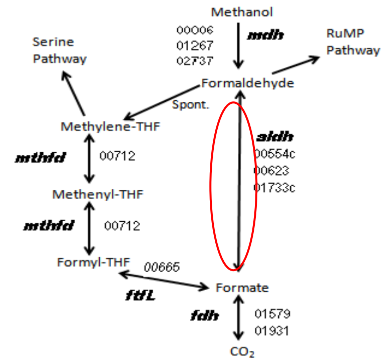


Figure 5.12. Regulation of putative aldehyde dehydrogenase (ORF 00623) in response to formaldehyde (2 mM), formate (2 mM) and methanol (for mannitol-grown cells, 100 mM). *B. methanolicus* MGA3 was grown in either methanol (Met) or mannitol (Man) and spiked when the O.D_{600} is ~0.8. The total RNA was extracted 20 minutes post-spiking.Met; immediately before spike, Met-FA; spiked with formaldehyde, Met-Formate; spiked with formate, Man; immediately before spike, Man-FA; spiked with formaldehyde, Man-Formate; spiked with formate, Man-Met: spiked with methanol. (n=3)



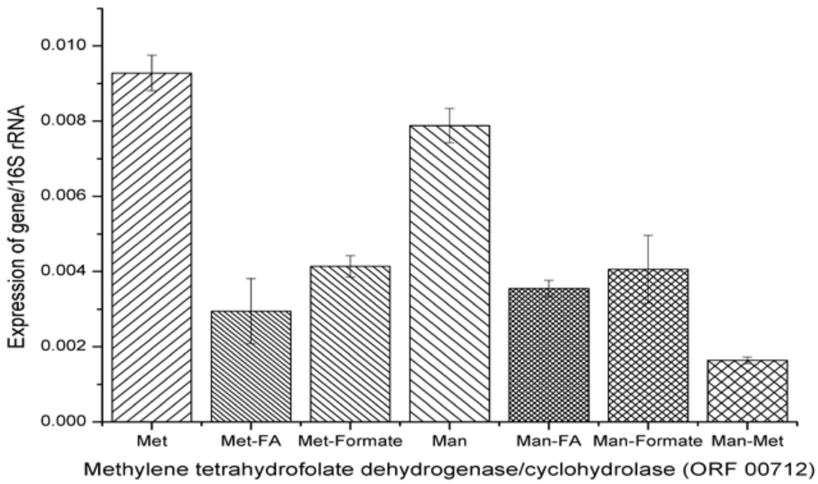
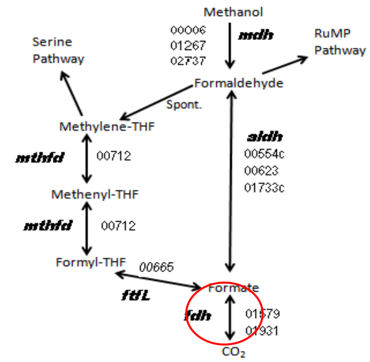


Figure 5.13. Regulation of putative methylene tetrahydrofolate dehydrogenase/cyclohydrolase (ORF 00712) in response to formaldehyde (2 mM), formate (2 mM), and methanol (100 mM) spikes. *B. methanolicus* MGA3 was grown in either methanol (Met) or mannitol (Man) and spiked when the $O.D_{600}$ is ~0.8. The total RNA was extracted 20 minutes post-spiking.Met; immediately before spike, Met-FA; spiked with formaldehyde, Met-Formate; spiked with formate, Man; immediately before spike, Man-FA; spiked with formaldehyde, Man-Formate; spiked with formate, Man-Met: spiked with methanol. (n=3)



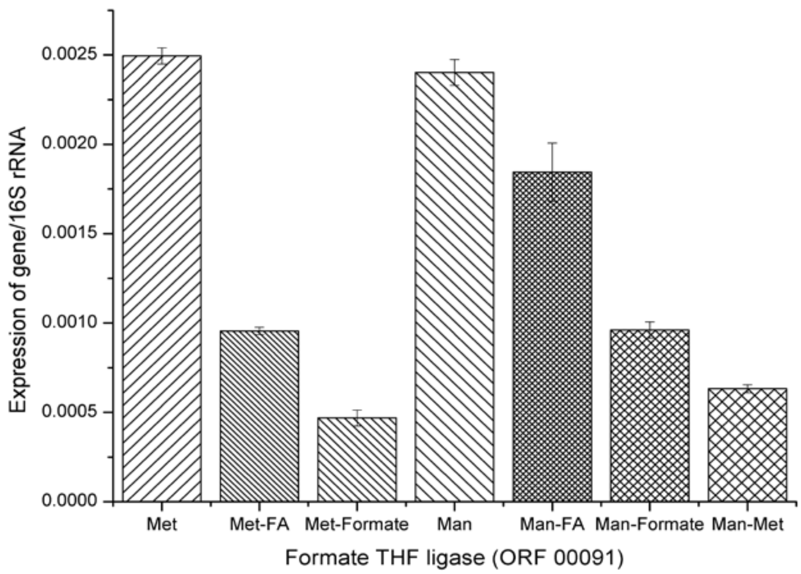
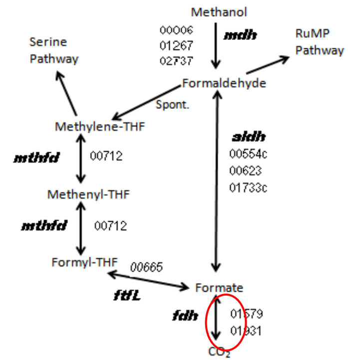


Figure 5.14. Regulation of putative formate THF ligase gene (ORF 00665) in response to formaldehyde (2 mM), formate (2 mM), and methanol (100 mM) spikes.*B. methanolicus* MGA3 was grown in either methanol (Met) or mannitol (Man) and spiked when the O.D$_{600}$ is ~0.8. The total RNA was extracted 20 minutes post-spiking.Met; immediately before spike, Met-FA; spiked with formaldehyde, Met-Formate; spiked with formate, Man; immediately before spike, Man-FA; spiked with formaldehyde, Man-Formate; spiked with formate, Man-Met: spiked with methanol. (n=3)



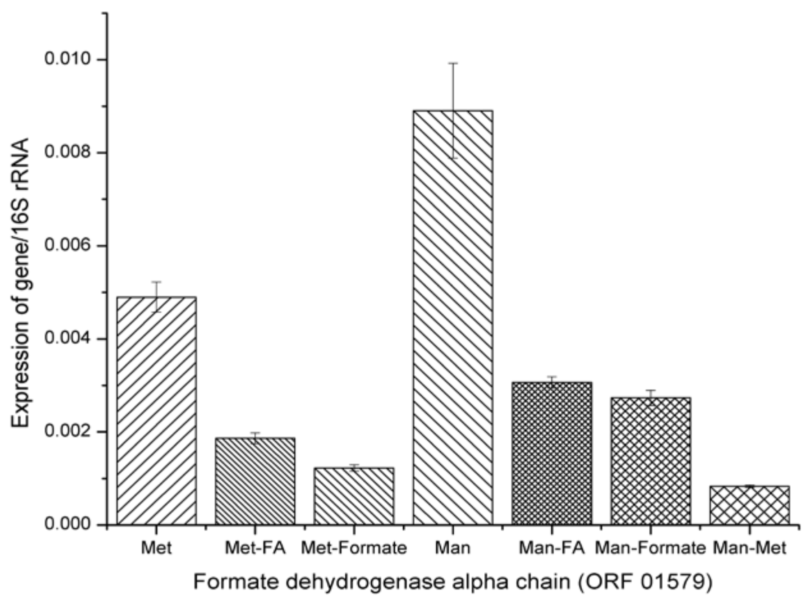
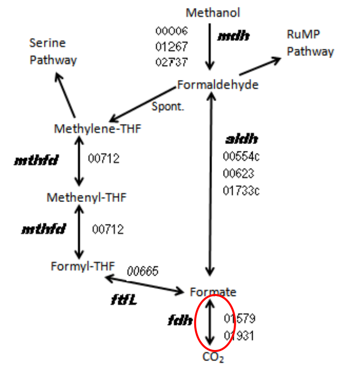


Figure 5.15. Regulation of putative formate dehydrogenase gene (ORF 01579) in response to formaldehyde (2 mM), formate (2 mM) and methanol (100 mM). *B. methanolicus* MGA3 was grown in either methanol (Met) or mannitol (Man) and spiked when the $O.D_{600}$ is ~0.8. The total RNA was extracted 20 minutes post-spiking.Met; immediately before spike, Met-FA; spiked with formaldehyde, Met-Formate; spiked with formate, Man; immediately before spike, Man-FA; spiked with formaldehyde, Man-Formate; spiked with formate, Man-Met: spiked with methanol. (n=3)



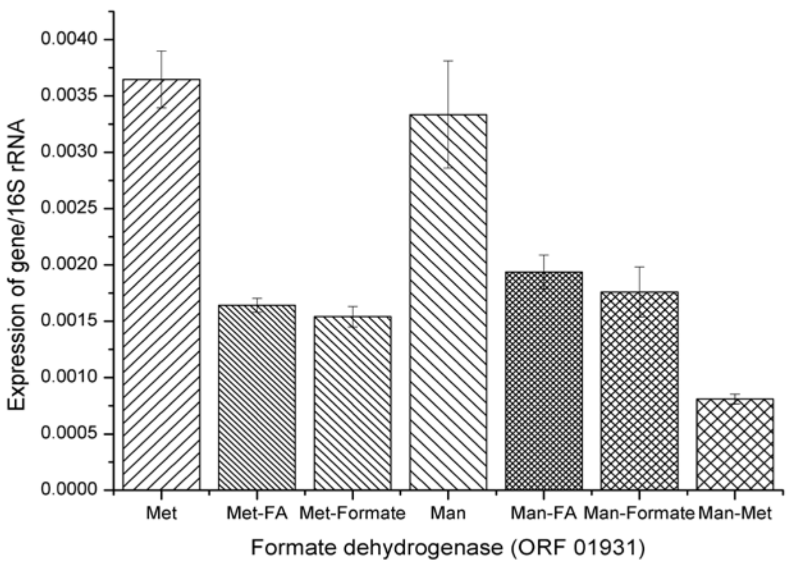
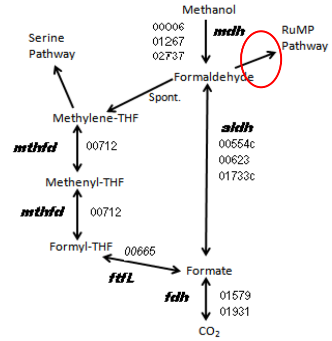


Figure 5.16. Regulation of putative formate dehydrogenase gene (ORF 01931) in response to formaldehyde (2 mM), formate (2 mM) and methanol (100 mM). *B. methanolicus* MGA3 was grown in either methanol (Met) or mannitol (Man) and spiked when the $O.D_{600}$ is ~0.8. The total RNA was extracted 20 minutes post-spiking.Met; immediately before spike, Met-FA; spiked with formaldehyde, Met-Formate; spiked with formate, Man; immediately before spike, Man-FA; spiked with formaldehyde, Man-Formate; spiked with formate, Man-Met: spiked with methanol. (n=3)



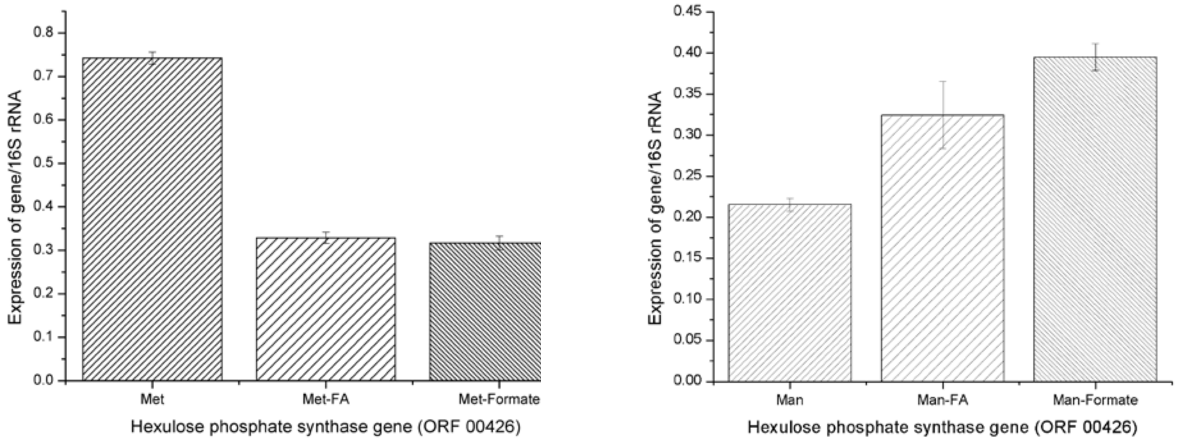
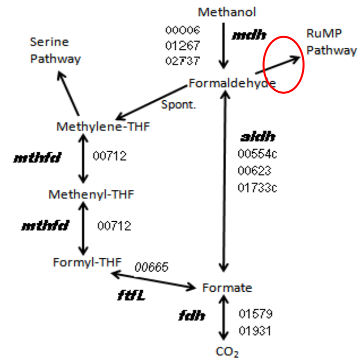


Figure 5.17. Regulation of hexulose phosphate synthase (*hps*) in response to formaldehyde (2 mM), formate (2 mM) and methanol (100 mM) spikes in methanol and mannitol-grown cells. *B. methanolicus* MGA3 was grown in either methanol (Met) or mannitol (Man) and spiked when the $O.D_{600}$ is ~0.8. The total RNA was extracted 20 minutes post-spiking.Met; immediately before spike, Met-FA; spiked with formaldehyde, Met-Formate; spiked with formate, Man; immediately before spike, Man-FA; spiked with formaldehyde, Man-Formate; spiked with formate. (n=3)



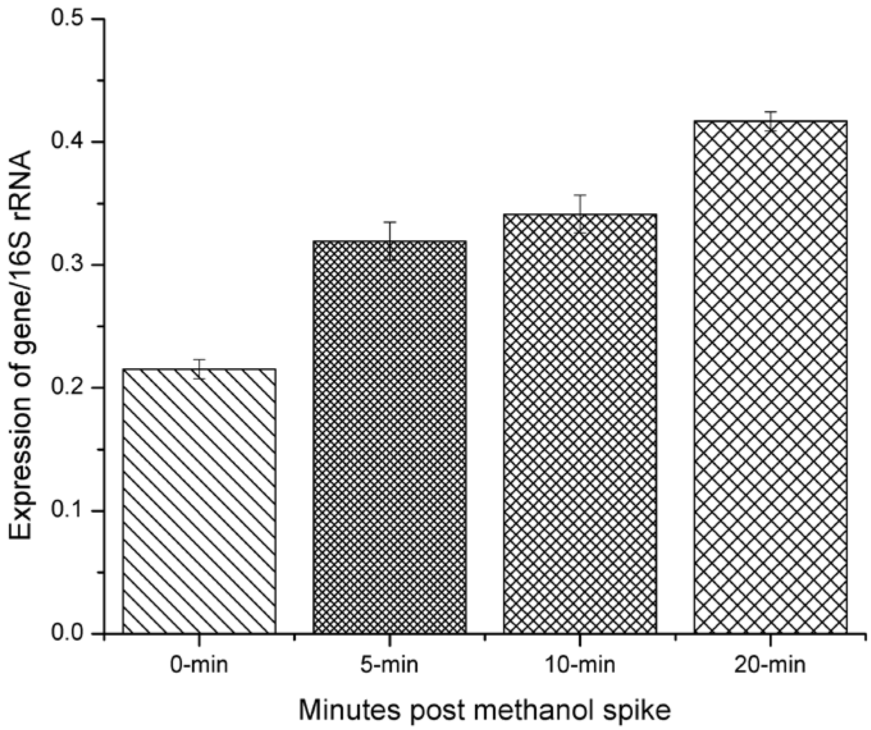


Figure 5.18. Regulation of hexulose phosphate synthase (*hps*) in response to methanol (100 mM) spike.*B. methanolicus* MGA3 was grown in $MTYM_{MAN}$ and spiked with methanol at $O.D_{600}$ of ~0.8. Then, the total RNA was extracted as described after 5, 10, and 20 min. The sample "0-min" was taken immediately before the spike. (n=3)

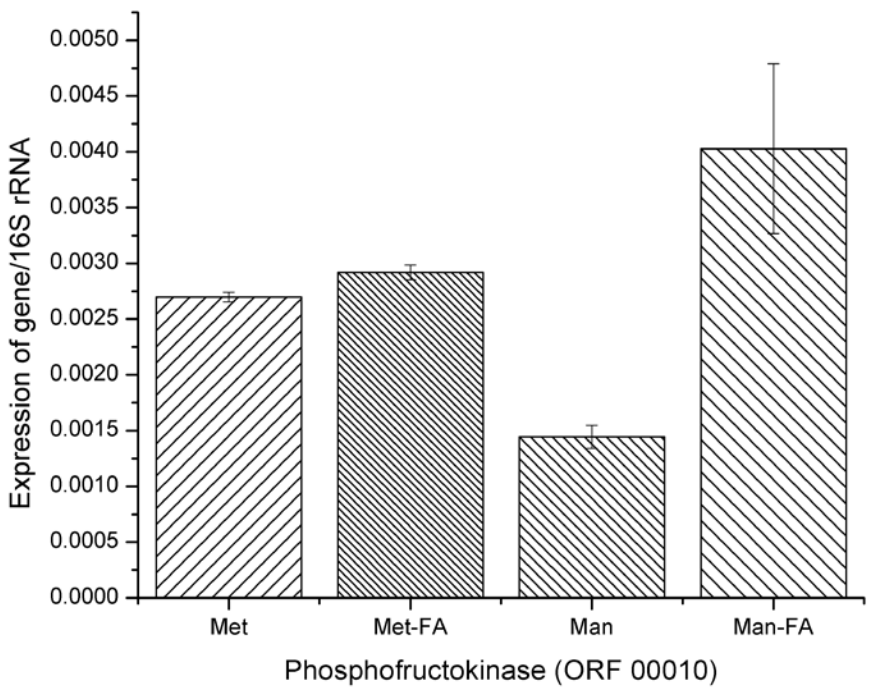


Figure 5.19. Regulation of phosphofructokinase (*pfk*) in response to formaldehyde (2 mM) spike.*B. methanolicus* MGA3 was grown in either methanol (Met) or mannitol (Man) and spiked when the $O.D_{600}$ is ~0.8. The total RNA was extracted 20 minutes post-spiking.Met; immediately before spike, Met-FA; spiked with formaldehyde, Man; immediately before spike, Man-FA; spiked with formaldehyde. (n=3)

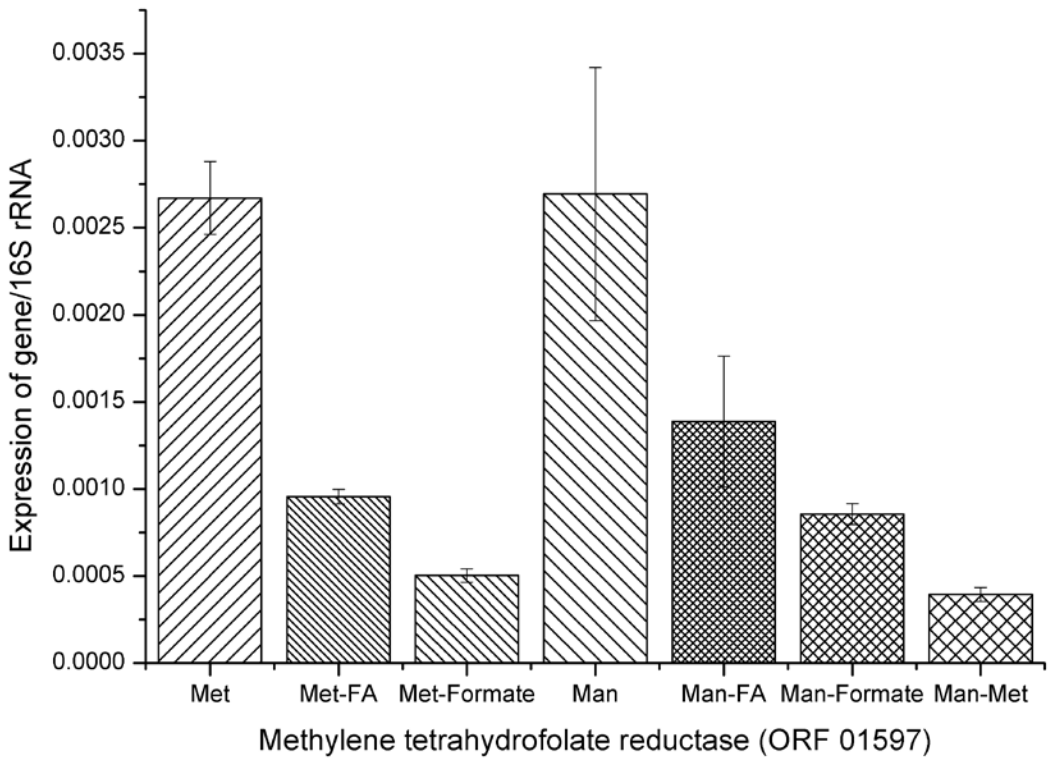


Figure 5.20. Regulation of methylene tetrahydrofolate reductase in response to formaldehyde (2 mM), formate (2 mM) and methanol (100 mM).*B. methanolicus* MGA3 was grown in either methanol (Met) or mannitol (Man) and spiked when the $O.D_{600}$ is ~0.8. The total RNA was extracted 20 minutes post-spiking.Met; immediately before spike, Met-FA; spiked with formaldehyde, Met-Formate; spiked with formate, Man; immediately before spike, Man-FA; spiked with formaldehyde, Man-Formate; spiked with formate, Man-Met: spiked with methanol. (n=3)

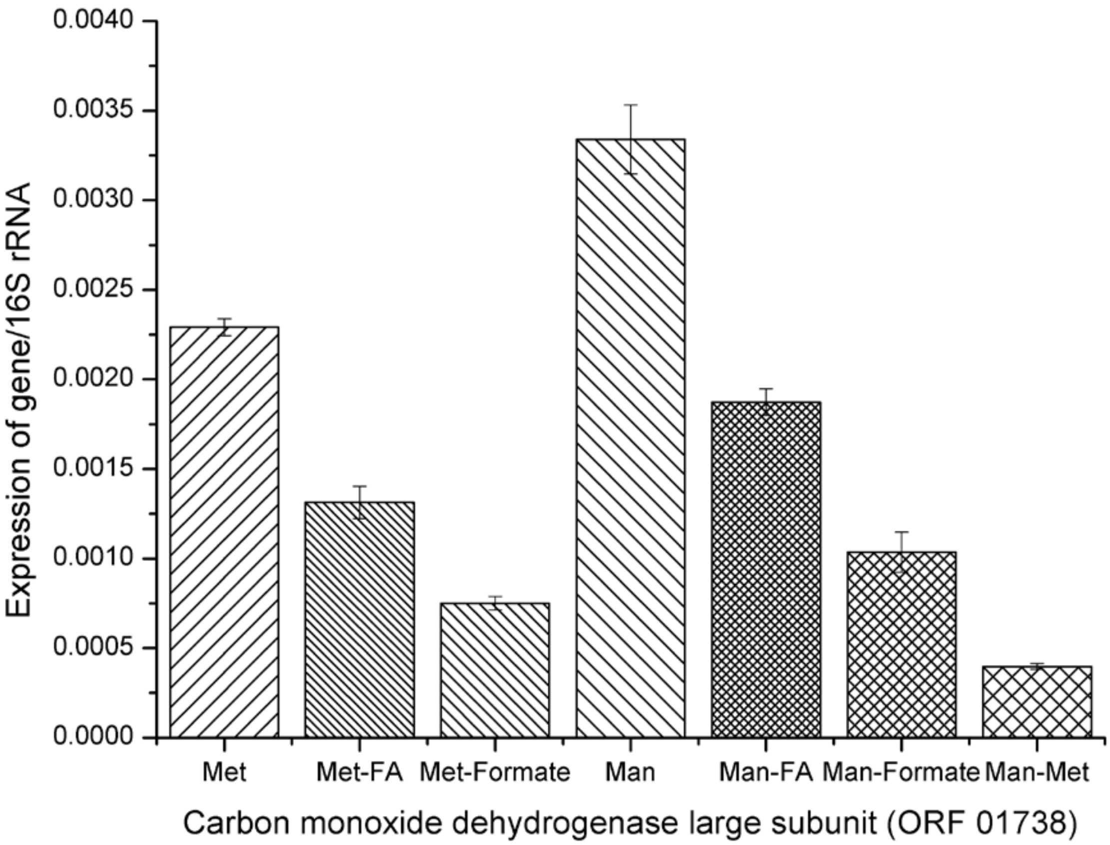


Figure 5.21. Regulation of CO dehydrogenase in response to formaldehyde (2 mM), formate (2 mM) and methanol (100 mM).*B. methanolicus* MGA3 was grown in either methanol (Met) or mannitol (Man) and spiked when the O.D_{600} is ~0.8. The total RNA was extracted 20 minutes post-spiking.Met; immediately before spike, Met-FA; spiked with formaldehyde, Met-Formate; spiked with formate, Man; immediately before spike, Man-FA; spiked with formaldehyde, Man-Formate; spiked with formate, Man-Met: spiked with methanol. (n=3)

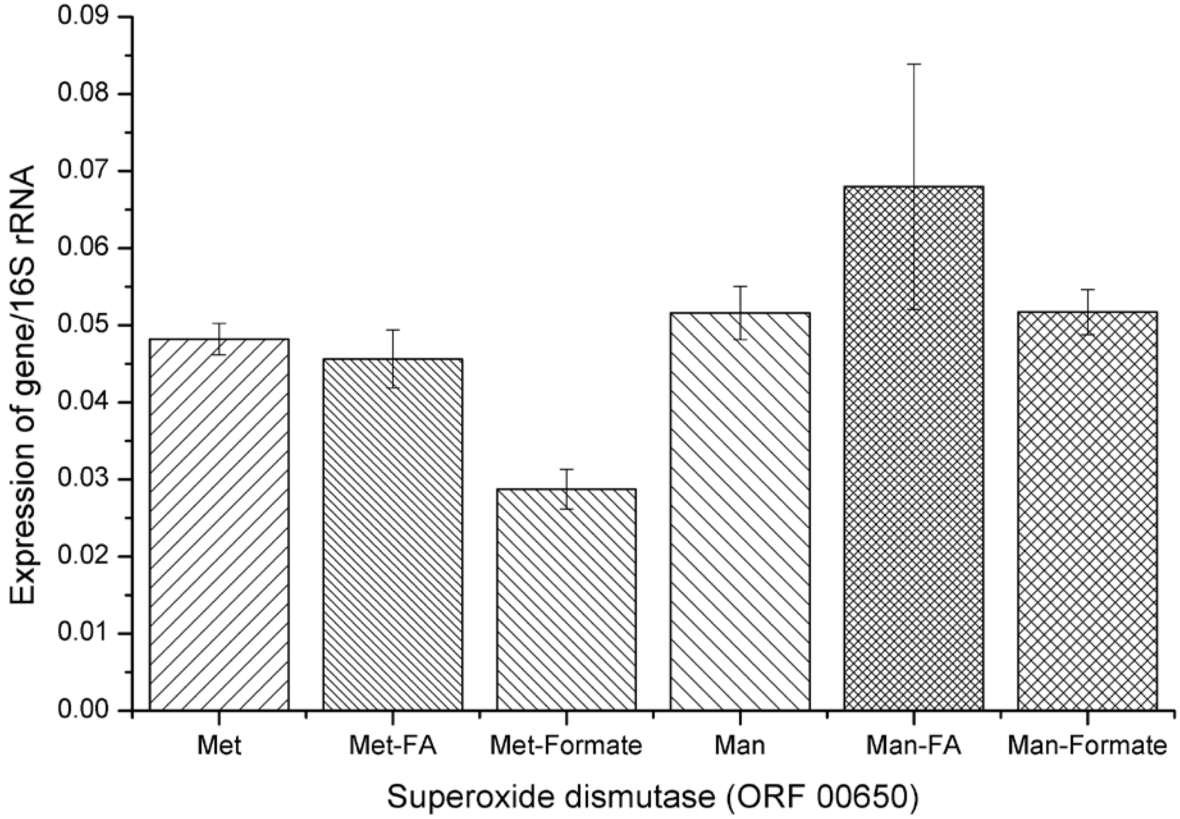


Figure 5.22. Regulation of superoxide dismutase in response to formaldehyde (2 mM), formate (2 mM) and methanol (100 mM).*B. methanolicus* MGA3 was grown in either methanol (Met) or mannitol (Man) and spiked when the $O.D_{600}$ is ~0.8. The total RNA was extracted 20 minutes post-spiking.Met; immediately before spike, Met-FA; spiked with formaldehyde, Met-Formate; spiked with formate, Man; immediately before spike, Man-FA; spiked with formaldehyde, Man-Formate; spiked with formate, Man-Met: spiked with methanol. (n=3)

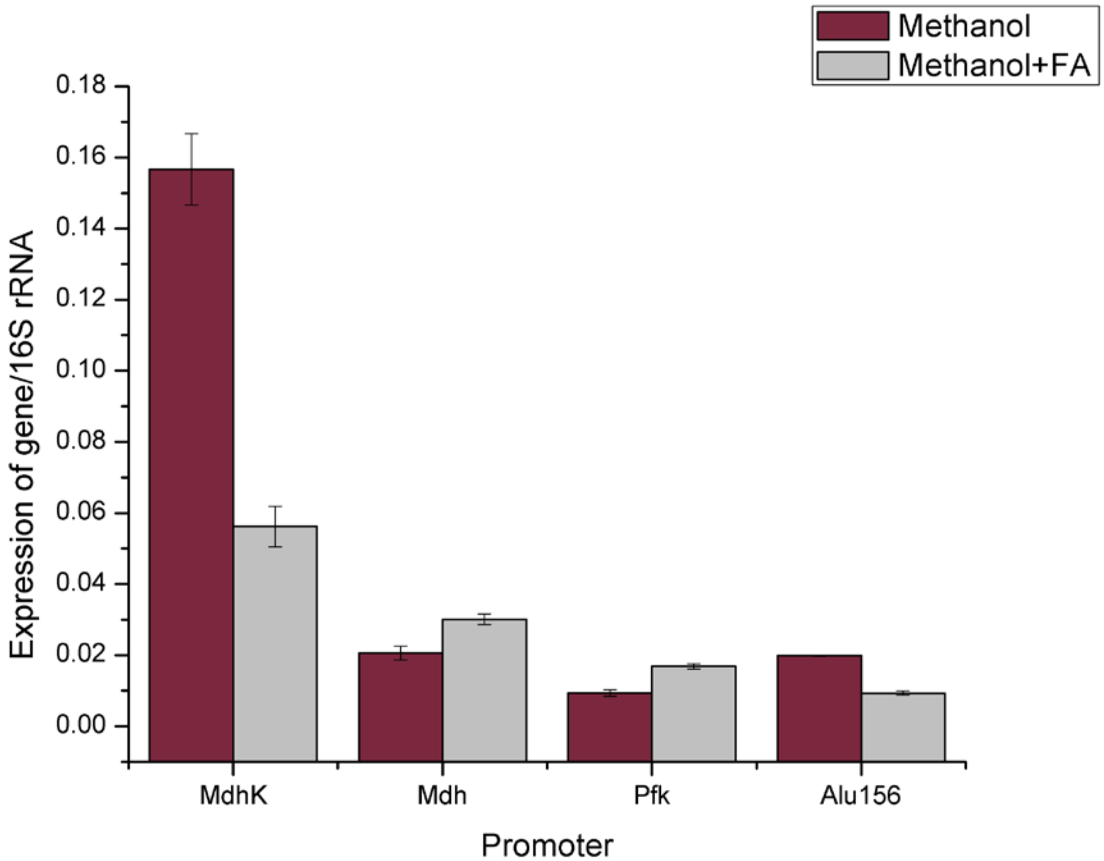


Figure 5.23. Expression GFPuv under different promoters.*B. methanolicus* MGA3 was grown in $MTYM_{150}$ (sample Methanol) and spiked with formaldehyde (sample Methanol+FA) at OD_{600} of ~0.8. The total RNA was extracted 20 min post-spiking. MdhK; GFPuv expressed from pNW33N as described in Nilasari et. al., (20012), Mdh; GFPuv expressed under P_{mdh} from pHP13, Pfk:GFPuv expressed under P_{pfk} from pHP13,Alu156: GFPuv expressed under P_{Alu156} from pHP13. (n=3)

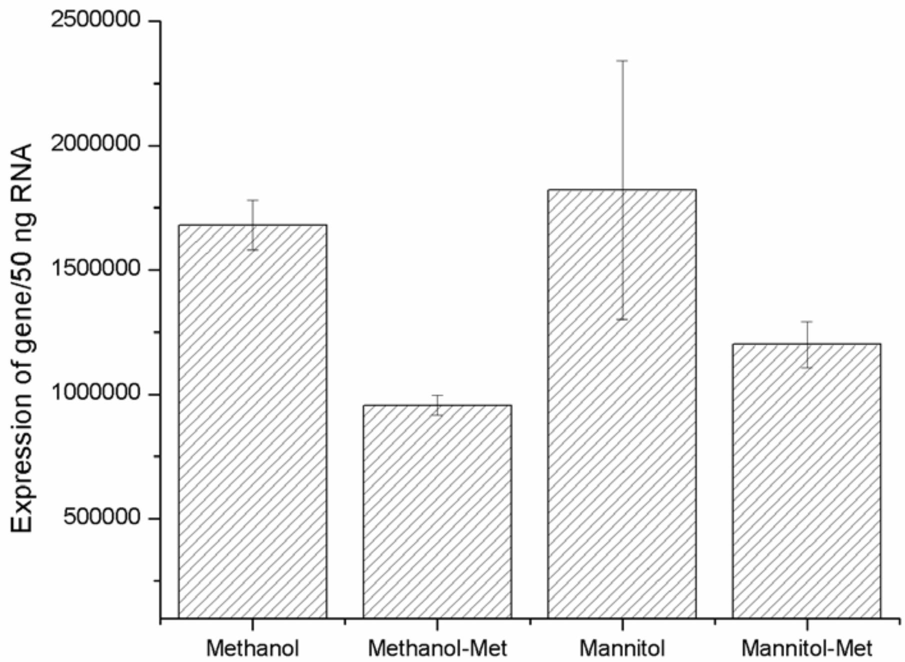


Figure 5.24. The expression level of *repB* in *B. methanolicus* MGA3. Met: Methanol grown cells, Met-Met: Methanol grown cells spiked with methanol, Man: Mannitol grown cells, and Man-Met: Mannitol grown cells spiked with methanol. Fifty ng of total rRNA was used for the assay. (n=3)

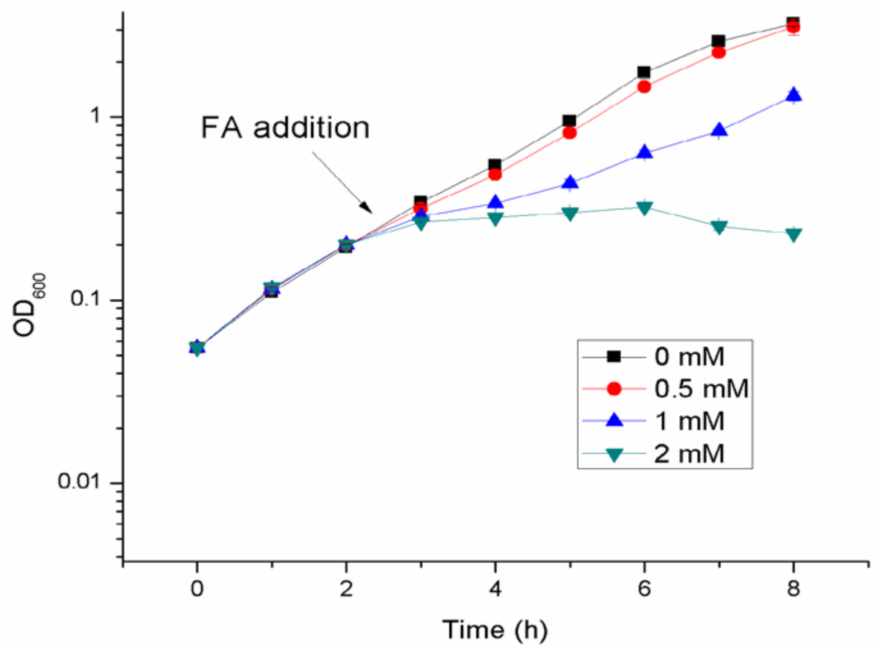


Figure 5.25. Formaldehyde tolerance of recombinant *B. methanolicus*MGA3(pTS123). *B. methanolicus* MGA3 (pTS123) was grown in MTYM with 150 mM methanol and spiked with formaldehyde at OD_{600} of ~0.250. (n=3)

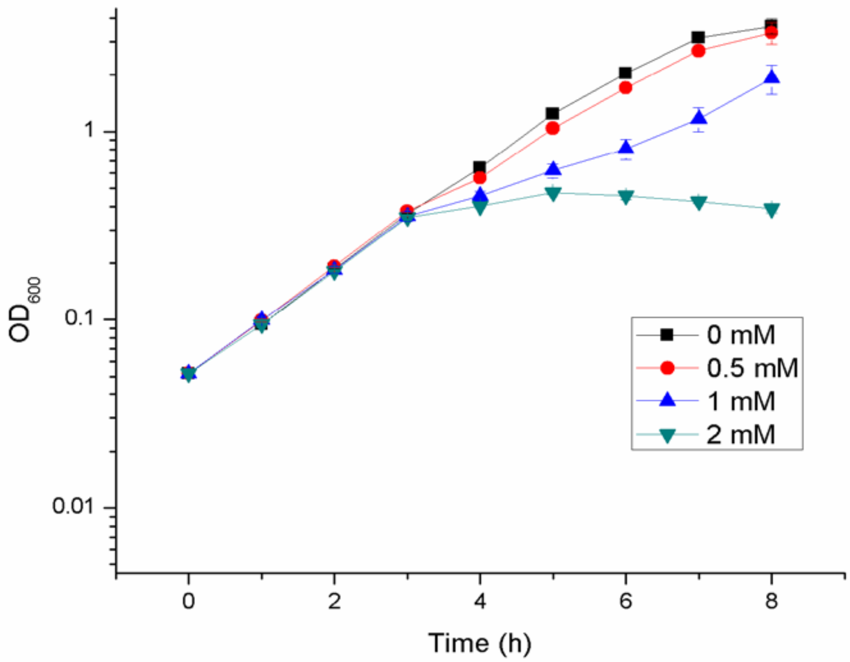


Figure 5.26. Formaldehyde tolerance of recombinant *B. methanolicus*MGA3(pHPS12). *B. methanolicus* MGA3 (pHPS12) was grown in MTYM with 150 mM methanol and spiked with formaldehyde at OD_{600} of ~0.250. (n=3)

CHAPTER 6

FORMALDEHYDE AND FORMATE DEHYDROGENASES OF *Bacilli*

Introduction

The concentration of CO_2in the atmosphere has been increasing since the start of the Industrial age, mostly due to the use of fossil fuels. Since the start of the Industrial Revolution, the carbon concentration has risen 30% (Vitousek et. al., 1997). This accumulation of CO_2 causes irreversible effects on the climate that last for 1000 years (Solomon et. al, 2009).

There are constant efforts for new technologies to limit the release of CO_2 into the atmosphere and fixation of atmospheric carbon to use as fuel, and as a substrate for production of intermediate products. One way of limiting the CO_2 emission is trapping it at the end of a pipeline. This is usually achieved by dissolving CO_2 in aqueous solution and absorbing it to an absorber like NaOH or seawater (Lin et.al. 2003). The second method of sequestering CO_2 is achieved through biology, which is conducted by nature; fixing it into biomass. For this, $CO_2$from the atmosphere is assimilatedinto the biomass, and biomass is then used as a substrate for production of biofuel and other metabolic intermediates. In this regards, the anaerobic bacterium *Clostridium ljungdahlii* was used to produce bioethanol using synthesis gas as substrate (Tirado-Acevedo, 2010).

Yet, another possibility is the enzymatic reduction of CO_2 to methanol, instead of ethanol. This pathway employs aerobic methylotrophic enzymes to synthesize methanol from CO_2. Practically, this is the reverse of the pathway that methylotrophic microorganisms use to dissimilate methanol to CO_2as discussed in Chapter 1. The simplest model of this pathway includes methanol dehydrogenase, formaldehyde dehydrogenase and formate dehydrogenase. Theoretically, formate dehydrogenase reduces CO_2 to formate which is then reduced to formaldehyde by formaldehyde dehydrogenase. Finally, formaldehyde is reduced to methanol by methanol dehydrogenase. Baskaya et. al. (2010), reported the thermodynamic feasibility of this route. The reverse of this biological pathway is feasible at low pH, low ionic strength and elevated temperatures. The biggest pitfall in this pathway is the source of reducing equivalentsin the form of NADH.

The enzymes of this route were discussed in Chapter 1. As a thermotolerant methylotrophic microorganism, *B. methanolicus* might serve as a source for these enzymes. The growth of *B. methanolicus* PB1 in sea-water based media at 50 °C was shown by Komives et. al. (2005).*B. methanolicus* carries methanol dehydrogenase genes both on a plasmid and on the chromosome. The presence of putative aldehyde (formaldehyde) and formate dehydrogenases was presented in Chapter 3. The methanol dehydrogenase enzyme from*B. methanolicus* PB1 was characterized before, although it is not identical to methanol dehydrogenase genes of strain MGA3 carried both on the pBM19 plasmid and chromosome (See Figure 3.4).

In this chapter, I report cloning, expression and characterization of putative formaldehyde and formate dehydrogenase genes of *B. methanolicus* MGA3, along with the putative respective genes of *Bacillus subtilis* BD170, *Bacillus licheniformis* and *Geobacillus kaustophilus*.

Materials and Methods:

Growth of microorganisms: *Bacillus subtilis*, *Bacillus licheniformis*, and *Geobacillus kaustophilus* HTA426 were obtained from ATCC. They were grown in 250 ml baffled shake flasks fitted with milk filter discs containing 50 ml of Luria-Bertani (LB) with 270 rpm shaking. The growth temperatures for microorganisms were 37 °C for *B. subtilis*, and *B. licheniformis*, and 60 °C for *Geobacillus kaustophilus* HTA426. *B. methanolicus* MGA3 was grown as described in Chapter 2.

Cloning of genes: The *B. subtilis* genes *yycR* and *yrhE*, *B. licheniformis* genes *ywdH* and *ywrH*, *G. kaustophilus* ORFs between the nucleotides of 2797314-2798810 and 498622-501585 were amplified using the Expand High Fidelity PCR kit from Roche, and primers listed in Table 6.1. Correct band size was verified using agarose gel electrophoresis. The PCR mixture was purified using Wizard® SV Gel and PCR Clean-Up System (Promega cat. A9281) and ligated into cloning vector pGEM-T using T_4 ligase (New England Biolabs, Ipswich MA). The ligated pGEM-T was transformed into *E. coli* DH5α using $CaCl_2$ competency as described in Maniatis et.al. (1982). Colonies were screened for the insert using PCR, and pGEM-T was isolated using Qiagen's QIAprep Spin Miniprep Kit. The insert was then subcloned into pET16-b or pET17-b using the respective restriction sites as listed in Table 6.1. The ligation, transformation and screening were carried out again as above and presence of correct size insert was verified using restriction enzyme digestion. For *B. methanolicus* MGA3+pTS123, the putative formaldehyde dehydrogenase gene of *B. kaustophilus* (2797314-2798810) was cloned into pHP13 under the control of P_{mdh} promoter. The manipulations were carried out using *E. coli* and then transformed into *B. methanolicus*

MGA3 as described in Chapter 5. The expression vectors were then submitted for sequencing for verification of nucleotide sequence. The same pattern was followed for the cloning of ORFs from *B. methanolicus* MGA3 except for subcloning into pGEM-T. The PCR amplified ORFs were directly cloned into either pET16-b or pET17-b, and nucleotide composition was verified again via sequencing.

Expression of recombinant genes: The expression of genes was carried out in *E. coli* BL21 (DE3). The host *E. coli* was grown in the presence of ampicillin at 37 °C, in 25 ml LB media in 250 ml shake-flasks to OD_{600} of ~0.8 and induced by 1 mM final IPTG. The temperature was lowered to 30 °C and culture was incubated for four hours. Cells were harvested by centrifugation for 5 min at 10,000xg and washed with 10 ml phosphate buffer saline (PBS). The pellet was resuspended in 2 ml PBS.

Lysis of cells: Twenty-microliter of Halt Protease Inhibitor Cocktail was added to resuspended cells in PBS and the suspension was transferred to 15 ml falcon tubes. Cell lysis was achieved using a Qsonica sonicator with micro tip in ice. The sonication parameters were; 50% probe density, 10 sec on, 20 sec off for one and half minute for *E.coli* and 3 minutes for *B. methanolicus*. The lysis of *B. methanolicus* differed from *E. coli* in that exponentially growing cells (OD_{600} ~1.5-2.0) were harvested and resuspended in 5 ml PBS with 20 mg/ml lysozyme after washing in 10 ml PBS. The cells were then incubated at 37 °C for an hour to digest the cell wall, washed again with 10 ml PBS to remove remnants of lysozyme and resuspended in 2 ml PBS toa final OD_{600} of ~20. Twenty microliters of protease inhibitor cocktail were added before sonication. The lysate was transferred back to a 50 ml falcon tube and centrifuged for 10 min at 10,000xg to remove the cell debris. The supernatant was used as crude cell extract. The protein concentration was measured using a Bradford assay.

SDS-polyacrylamide gel electrophoresis (SDS-PAGE): SDS-PAGEwas performed using a Mini-Protean Tetra System (BioRad Laboratories). Fifteen microliters of sample were mixed with 15 µl 2x sample buffer containing 5% β-mercaptoethanol and denatured for 10 min at 90 °C before loading into a 12% nUView Tris-HEPES gel (NuSep, Bogart GA). The gel was run at 100 V for 45-60 minutes and stained in 25 ml Instant Blue (Kampenhout, Belgium). The gel was visualized using Kodak Gel Logic 200 imaging system after destaining with a solution containing 5% methanol and 7% acetic acid.

Purification of His-tagged proteins: Thermo Scientific HisPur Ni-NTA resin was used to purify the His-tagged proteins from the crude cell extracts according to the manufacturer's manual.

Formaldehyde dehydrogenase assay: The formaldehyde dehydrogenase of *Pseudomonas putida* (Sigma F1879, 3U/mg solid) was used as standard enzyme to construct a standard curve for the formaldehyde dehydrogenase activity. The entire sample in the tube was dissolved in 100 µl of PBS and protein content was measured using a Bradford assay. Then, it was diluted to the range of 0.01 to 0.0005 units, in PBS. Formaldehyde activity was assayed by measuring formation of diformazan. The reaction mixture, (1 ml) containing 26 mM potassium phosphate buffer (pH 8.0), 0.2% Triton X, 5 mM formaldehyde, 0.005 % NBT, 0.0005% PMS and 0.5 mM NAD^+, was prewarmed for 5 minutes in a water bath and the reaction was started by adding the enzyme in a 1 ml disposable cuvette. The change in OD_{570} was measured at exactly 3 minutes. The standard curve obtained is shown in Figure

6.1. One unit of enzyme is defined as amount of enzyme that will oxidize 1 µmol of formaldehyde into formic acid per minute.

Formate dehydrogenase assay: The assay for formate dehydrogenase was the same as with formaldehyde dehydrogenase except that formaldehyde was replaced with formate. The standard curve obtained is shown in Figure 6.2. A standard enzyme isolated from*Candida boidinii* was purchased from Sigma. One unit of enzyme is defined as amount of enzyme required to oxidize 1 µM of formate into CO_2 per min.

Results:

Expression of clones: The formaldehyde dehydrogenase genes of *B. subtilis* and *B. kaustophilus* were successfully expressed from pET16-b vector (Figure 6.3). However, the expression of formate dehydrogenase from *B. subtilis* (yrhE), *B. licheniformis* (yrhE) and *G. kaustophilus*(498622-501585) was not assessed. *B. methanolicus* MGA ORFs 01733 (putative formaldehyde dehydrogenase) and ORF 01931 (putative formate dehydrogenase) were successfully overexpressed in *E. coli* BL21 as seen in Figure 6.4. There was no distinct overexpression seen on the SDS-PAGE for clones of ORFs 00554, 00623 and 01579.

Formaldehyde dehydrogenase activity of *yycR*: The putative formaldehyde dehydrogenase of *B. subtilis* was successfully overexpressed and purified by metal affinity chromatography (Figure 6.4). Figure 6.5 describes the enzymatic assay of purified FADH of *B. subtilis* as a function of pH.

Formaldehyde dehydrogenase activities of *E. coli, B. subtilis, B. methanolicus* MGA3, *B. methanolicus* PB1 and *B. methanolicus* MGA3 (pTS123): The formaldehyde and formate dehydrogenase activities of samples from crude cell extracts were assayed using the calibration curves constructed via standard enzymes as described above. One unit of enzyme activity is defined as the amount of enzyme activity that oxidizes 1 µmol of substrate (formaldehyde or formate) into product per min. As seen from Table 6.2, *B. methanolicus* strains MGA3 and PB1 have significantly higher formaldehyde dehydrogenase activities than that of *B. subtilis* and *E. coli* BL21. Also, the formaldehyde dehydrogenase activity of *B. methanolicus* PB1 is two-fold higher than that of strain MGA3. This is consistent with Figure 1.4 which demonstrates a higher CO_2 dissimilation rate from the strain PB1 and also with the Sothern Blot data, on which the probes were able to hybridize with putative formaldehyde gene(s) of strain PB1 but not with that of MGA3. Although there was no evident expression on the gel (Figure 6.6), the difference between formaldehyde dehydrogenase activities of *E. coli* BL21 and *E. coli* BL21+ORF00554 and ORF00623 (Table 6.2) suggests possible expression of the respective genes, albeit very weakly. The specific formaldehyde dehydrogenase activity of *B. methanolicus* is comparable to that of *M. capsulatus* (Bath), (Adeosun et. al., 2004). This difference deserves further investigation. Additionally, no formaldehyde dehydrogenase was detected from*B. methanolicus* MGA3 (pTS123). This requires further investigation as*B. methanolicus* MGA3 (pTS123) is more tolerant to formaldehyde spike than the parent strain (Figure 5.25)

Formate dehydrogenase activities of *B. subtilis, E. coli, B. methanolicus* strains MGA3 and PB1 and ORF01931: Table 6.3 summarizes the formate dehydrogenase activity from crude cell extracts of the samples. The enzyme activities were measured as described above and calculated using the calibration curve in Figure 6.2. One unit of enzyme activity is defined as the activity that oxidizes 1 µmol of formate into CO_2 per minute. As seen in Table 6.3, the formate dehydrogenase activity of *B. methanolicus* strain PB1 was significantly higher than that of strain MGA3. The formate dehydrogenase activities of *B. subtilis* and *E. coli* are comparable to each other. The possible overexpression of ORFs 01579 and 01931 (see figure 6.4 for overexpression of ORF01931) in *E. coli* caused two fold increase in specific activity of crude cell extract. In addition, the purified FDH (ORF1931) (Figure 6.7) had ~3-fold higher formate dehydrogenase activity than that of crude cell extract. However, this activity was too small for a purified enzyme. This might be either due to very low affinity of the expressed protein, the His-tag that it contains may alter its activity or the reaction conditions may require further optimization.

Conclusion:

The putative formaldehyde and formate dehydrogenases of *B. subtilis, B. licheniformis, G. kaustophilus* HTA 426 and *B. methanolicus* MGA3 were investigated in this chapter. The crude cell extract of *B. subtilis*, and *B. methanolicus* strains MGA3 and PB1 were tested for formaldehyde dehydrogenase activity. Among them the crude cell extract of *B. methanolicus* PB1 had shown the highest enzymatic activity. Up to 7 mU/mg protein of activity was detected in *B. methanolicus* strain PB1 whereas the MGA3 strain had only half of it, 3.057 mU/mg protein. *B. subtilis* had even less formaldehyde dehydrogenase activity, 1.192 mU/mg protein. This activity of formaldehyde dehydrogenase from crude cell extracts of *B. methanolicus* is comparable to that of *M. capsulatus* Bath which had activity of 5 mU/mg protein (Adeosun et. al., 2004).

The difference in enzymatic activity was more profound between the formate dehydrogenases of strains MGA3 and PB1, 3.053 mU/mg protein vs. 8.574 mU/mg protein. The activities of *B. subtilis* and *E. coli* BL21 are 1.473 and 1.873 mU/mg protein respectively. The *E. coli* strains that harbor putative *fdh* genes in pET system hada 2-fold increased formate dehydrogenase activity compared to the parent strain.

The simplest route for the dissimilation of methanol into CO_2 is oxidation through the enzymes of formaldehyde and formate dehydrogenases, which requires molecular O_2. Gutmacher (2009) has shown that the strain PB1 has higher O_2 demand during fed-batch fermentations (See figure 1.5), and incorporates less carbon into biomass per consumed methanol. The differences in carbon dissimilation pathway of strains MGA3 and PB1 were also emphasized in Chapter 4 with Southern Blot. Since the genome sequence revealed putative formate dehydrogenases in both strains (Chapter 2 and Heggeset et al., 2012), the difference in CO_2 evolution rate was most probably due to the efficiency of these enzymes.

The purified FDH of *B. methanolicus* MGA3 had activity of 13.116 mU/mg protein. This is a significant increase from the crude cell extract, but further investigation is required to optimize the purification and enzymatic assay for FDH. It is possible that the reaction conditions or expression parameters require optimization for maximum activity. Also, the FDH of strain PB1 may have a better activity, given the crude cell extract of *B. methanolicus* PB1 is 3-fold higher than the strain MGA3.

Among the FADHs of *B. subtilis, B. licheniformis* and *G. kaustophilus,* the *B. subtilis* and *G. kaustophilus* enzymes were overexpressed in *E. coli* as seen in Figure 6.3. The *B. subtilis* FADH was further characterized and compared to standard enzyme from *P. putida* (Degenkolbe, 2010). The activity of *B. subtilis* FADH with varying temperatures is summarized in Figure 6.5. Despite the initial demonstration of FADH activity of *G. kaustophilus* at 65 °C (Figure 6.9), no activity was detected when the crude cell extract was tested at 37 °C. The optimal growth temperature of *G.kaustophilus* is 65 °C and this might be the reason for lacking activity at 37 °C. This clone also needs further investigation.

Table 6.1. Primers used to amplify the respective gene for cloning and target cloning vectors. The restriction sites for respective enzymes are underlined.

Microorganism	Gene	Primer	Exp. Vector	Sequence
B. subtilis	*yycR*	Bs_faldh_Fw	pET16-b	5'-ATG GTA CTC GAG ATG GGA GGA ATG GCG TTG ACA G-3'(XhoI)
B. subtilis	yycR	Bs_faldh_Rv	pET16-b	5'-ATG GTA GGA TCC TTA CTT CAA TGT CCC ATG-3' (BamHI)
B. subtilis	yrhE	Bs_fdh_Fw	pET16-b	5'-ATG GTA CTC GAG ATG AGA TCT TTT GCC ACA CAA C- 3'(XhoI)
B. subtilis	yrhE	Bs_fdh_Rv	pET16-b	5'-ATG GTA GGA TCC TCA ATC TGT ACA CGG CCT G-3' (BamHI)
B. licheniformis	ywdH	Bl_faldh_Fw	pET16-b	5'- ATG GTA CAT ATG GCA GAA CGA ACA ACC TCC-3' (NdeI)
B. licheniformis	ywdH	Bl_faldh_Rv	pET16-b	5'-ATG GTA GGA TCC CTA CCT GAG AAG TTT TTT G-3' (BamHI)
B. licheniformis	yrhE	Bl_fdh_Fw	pET16-b	5'- ATG GTA CAT ATG GAC GGA AAG CCA ATC AG-3' (NdeI)
B. licheniformis	yrhE	Bl_fdh_Rv	pET16-b	5'- ATG GTA GGA TCC TTA GTC TGT TAA ATG AAC G-3' (BamHI)
G. kaustophilus	2797314-	Gk_faldh_Fw	pET16-b	5'- ATG GTA CAT ATG ACA ATG ATT TCG GCG-3' (NdeI)
G. kaustophilus	2797314-	Gk_faldh_Rv	pET16-b	5'- ATG GTA GGA TCC TTA AAA CAA CGT CAT CGG-3' (BamHI)
G. kaustophilus	498622-	Gk_fdh_Fw	pET16-b	5'-ATG GTA CAT ATG GAA GCG CAA ATG GTG-3'(NdeI)
G. kaustophilus	498622-	Gk_fdh_Rv	pET16-b	5'-ATG GTA GGA TCC TTA GCC ACC GAT CGT CAC-3' (BamHI)
B. methanolicus	ORF00554	19-Alde Fw	pET17-b	5'-ATG GTA AAG CTTATGGAATATCAAGTTACG-3'
B. methanolicus	ORF00554	19-Alde Rv	pET17-b	5'-ATG GTA CTC GAGTTAAAACGGAAATTTGCG-3' (XhoI)
B. methanolicus	ORF00623	20-Alde Fw	pET16-b	5'-ATG GTA CAT ATGGTTGTTTTGAGTGAAACC-3' (NdeI)
B. methanolicus	ORF00623	20-Alde Rv	pET16-b	5'-ATG GTA GGA TCCTCAATGTAAGTTTACCC-3' (BamHI)
B. methanolicus	ORF01733	36 Alde-Fw	pET16-b	5'-ATG GTA CAT ATGAGTCATTTAACCGTCG-3'(NdeI)
B. methanolicus	ORF01733	36 Alde-Rv	pET16-b	5'-ATG GTA GGA TCCTTAGTTTAAAGAAATCC-3'(BamHI)
B. methanolicus	ORF01579	32-Form-Fw	pET17-b	ATG GTA AAG CTTATGGGAAAAACACAACATAC-3'
B. methanolicus	ORF01579	32-Form-Rv	pET17-b	ATG GTA CTC GAGTCGAGACGTAACATAAAAATGC-3'
B. methanolicus	ORF01931	48-Forma-	pET16-b	ATG GTA CTC GAGATGCACTCTTATGTGGATCAG-3'
B. methanolicus	ORF01931	48-Forma-	pET16-b	5'-ATG GTA GGA TCCTTACAGTTTCTCCACATC-3'

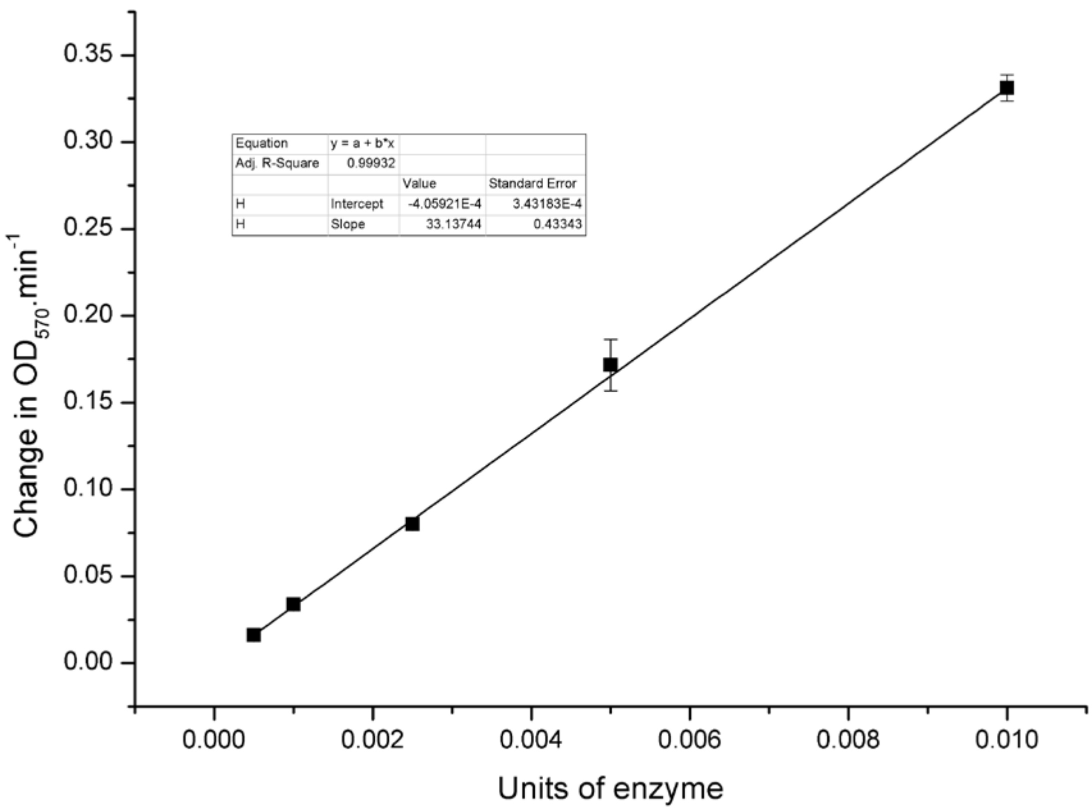


Figure 6.1. The calibration curve for measurement of formaldehyde dehydrogenase activity. *P. putida* formaldehyde dehydrogenase was used as the standard enzyme for the construction of standard curve.

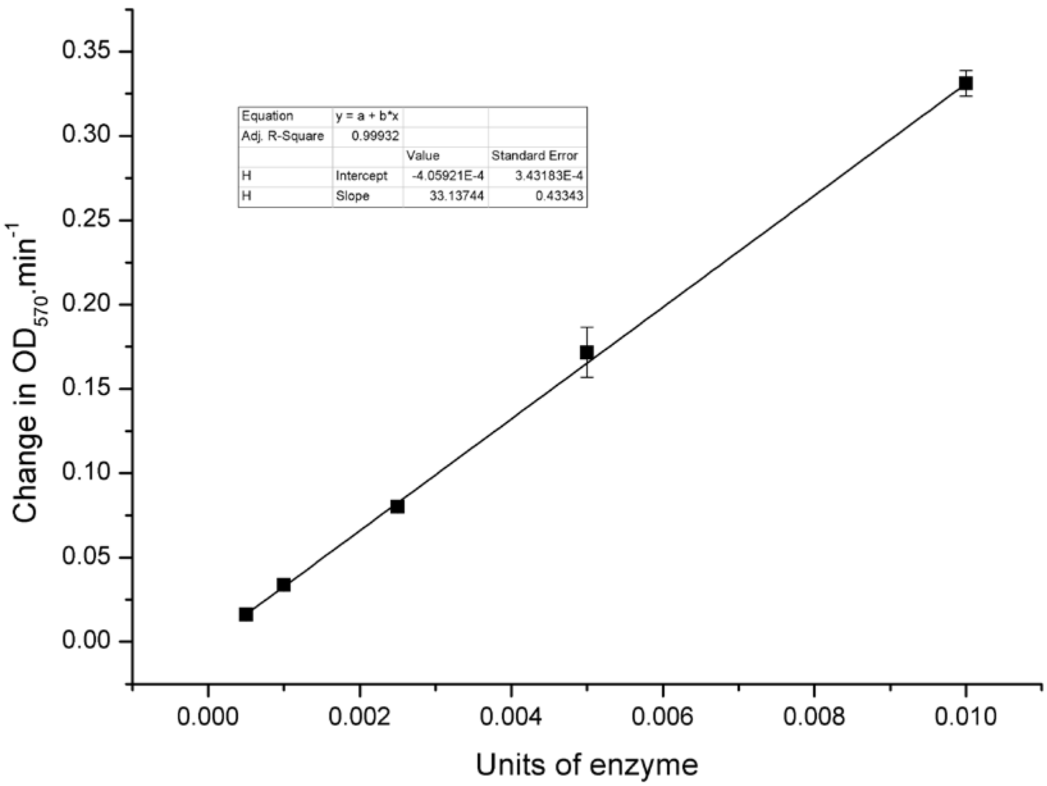


Figure 6.2. The calibration curve for measurement formate dehydrogenase activity. *Candida boidinii* formate dehydrogenase was used as standard enzyme for construction of standard curve.

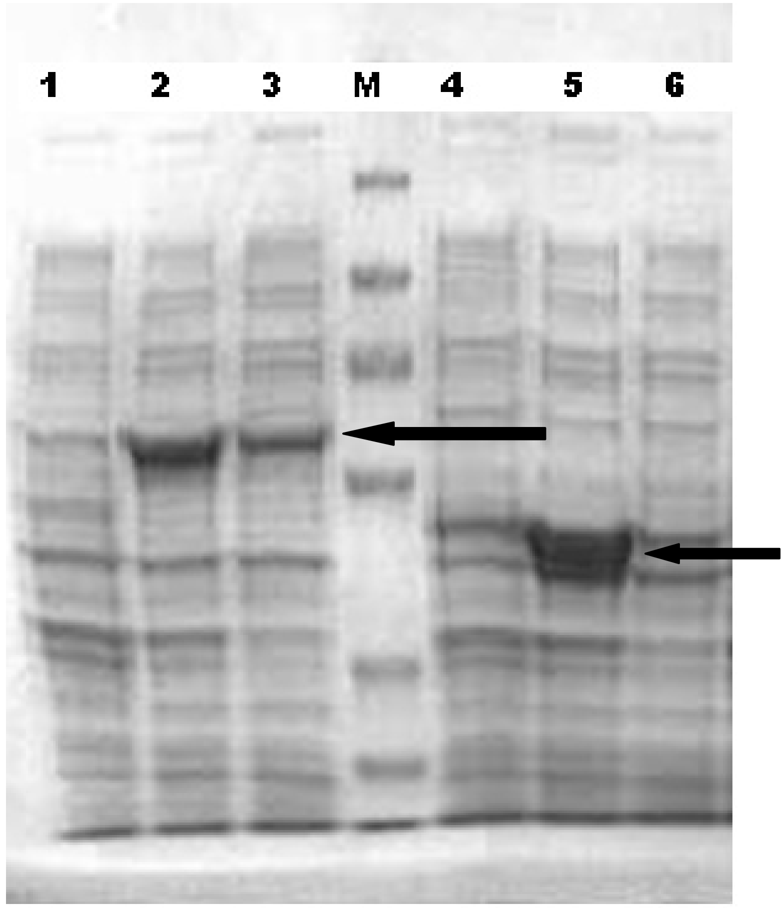


Figure 6.3. Overexpression of *B.kaustophilus* (53.49 kDa) and *B. subtilis faldh* (42.95 kDa) genes. Lines 1: *B. kaustophilus* uninduced, 2: *B. kaustophilus* induced insoluble, 3: *B. kaustophilus* induced soluble. M: Protein molecular marker, 4: *B. subtilis* uninduced, 5: *B. subtilis* induced insoluble, 6: *B. subtilis* soluble. (Image courtesy of Degenkolbe, 2010).

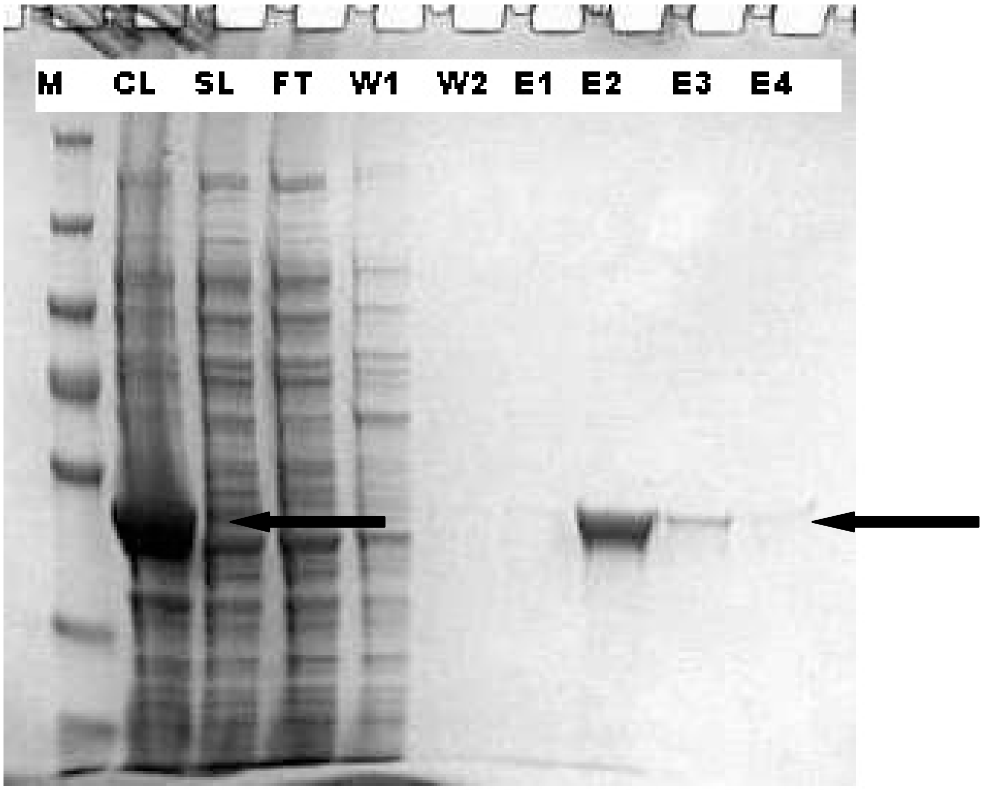


Figure 6.4. Purification of overexpressed *B. subtilis faldh* (42.95 kDa) via IMAC. M: Molecular marker, CL; crude cell lysate, SL; soluble fraction from the lysate, FT: flow through from IMAC column, W1 and W2: washes, E1-4: Elutions (Courtesy of Degenkolbe, 2010).

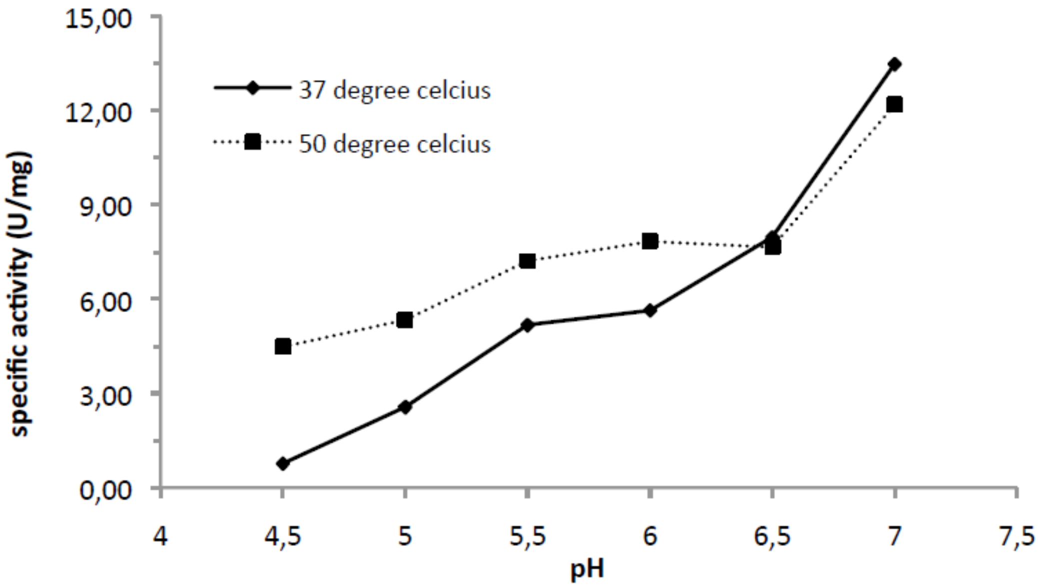


Figure 6.5. Specific activity of *B. subtilis* FADH at varying pH. (Degenkolbe, 2010)

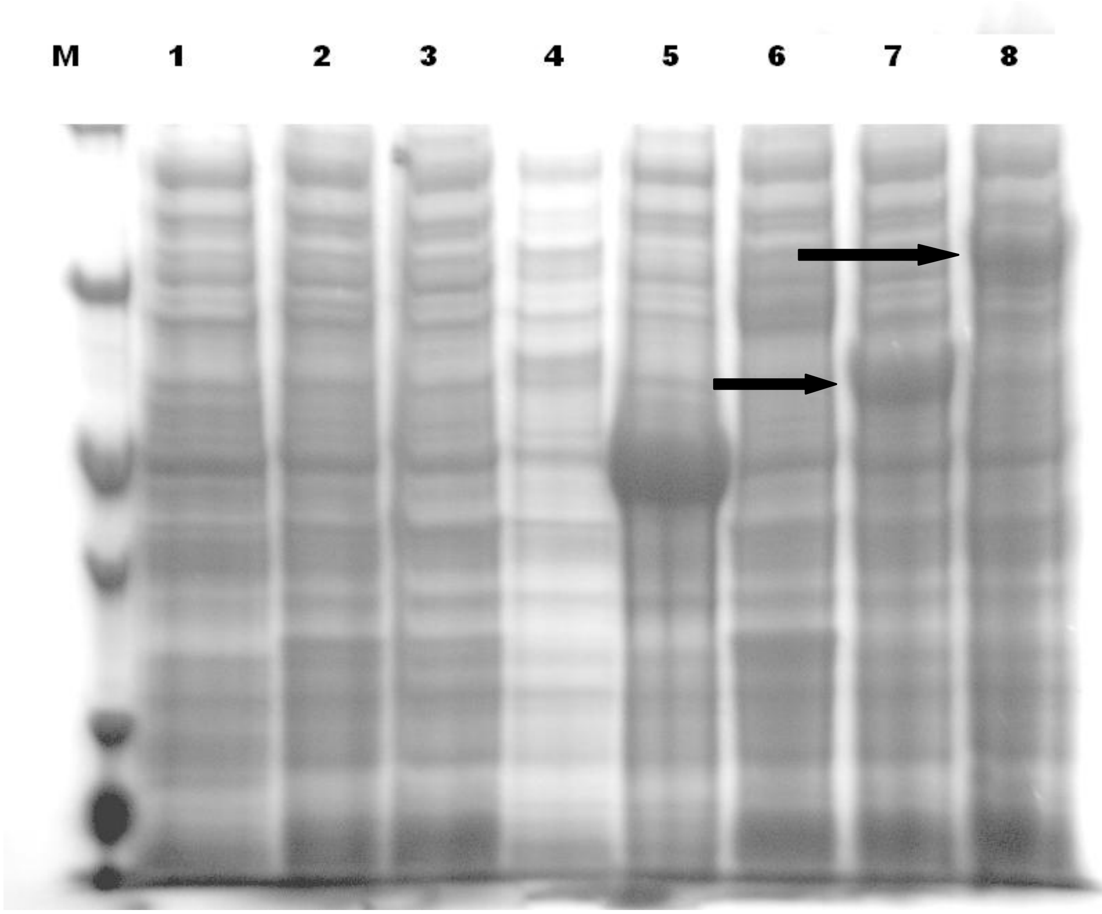


Figure 6.6. Analysis of crude cell extract of *E. coli* BL21 cells transformed with respective genes. M: marker, 1: *E. coli* BL21, 3: *E. coli* with ORF 554, 4: *E. coli* with ORF 623, 6: *E.coli* with ORF 1532, 7: *E. coli* with ORF 1733 (53.89 kDa) and 8: *E. coli* with ORF 1931 (75.69 kDa). The samples in lanes 2 and 5 are not directly related with this topic.

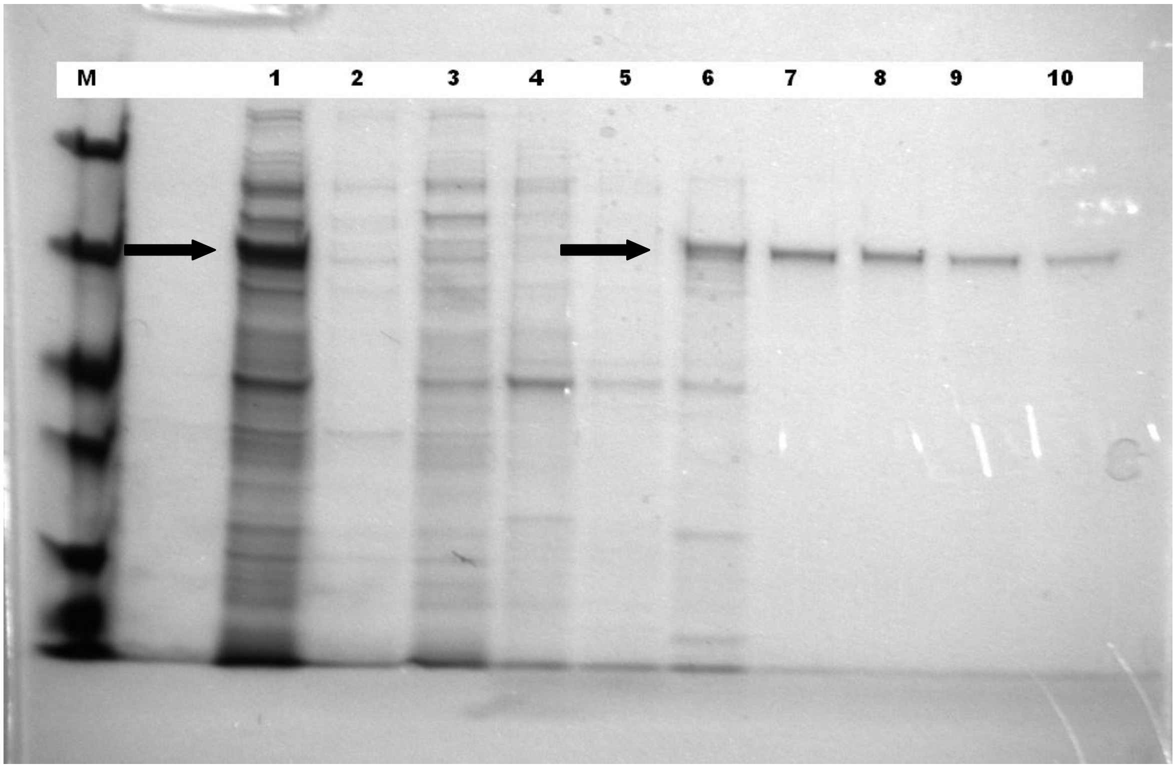


Figure 6.7. Purification of putative formate dehydrogenase (ORF01931-75.69 kDa) using IMAC. M: Molecular marker, 1: Crude cell extract, 2: Flow through, 3:Wash 1, 4: Wash 2, 5-10: Elution 1-6.

.

Table 6.2. Formaldehyde dehydrogenase activities from crude cell extract. One unit of enzyme oxidizes on µmol of formaldehyde into formate per min (n=3). ND: Not detected.

Sample	FADH activity mU/mg protein	STD
B. subtilis	1.19	±0.08
B. methanolicus MGA3	3.08	±0.12
B. methanolicus PB1	6.97	±1.94
B. methanolicus MGA3 pTS123	ND	ND
E. coli BL21	0.45	±0.02
E. coli BL21+ORF00554	3.76	±0.04
E. coli BL21+ORF00623	1.98	±0.03
E.coli BL21+ORF01733	0.86	±0.03

Table 6.3. Formate dehydrogenase activities of crude cell extract. One unit of enzyme oxidizes one µmol of formate into CO_2 per min.

Sample	FDH activity mU/mg protein	STD
B. subtilis	1.44	±0.06
B. methanolicus MGA3	3.05	±0.46
B. methanolicus PB1	8.57	±2.39
E. coli BL21	1.84	±0.05
E. coli BL21+ORF1579	3.84	±0.07
E. coli BL21+ORF1931	3.86	±0.17
Purified FDH (ORF 1931)	13.12	±1.46

CHAPTER 7

INVESTIGATION OF ANAEROBIC GROWTH OF *B. methanolicus* MGA3

Introduction

The anaerobic growth of a strict aerobe,*B. subtilis*, was discussed in Chapter 1. Glaser et. al., have grown *B. subtilis* anaerobically in rich medium (2XYT) supplemented with 0.2 % KNO_3 as terminal electron acceptor (Nakano and Zuber, 1998). Another bacillus, *Bacillus infernus* can grow on formate or lactate with Fe(III), MnO_2 or nitrate as an electron acceptor (Boone et. al., 1995). Interestingly, methylotrophic bacteria were also isolated from legume nodules, which are known for the requirement of strict anaerobic conditions (Sy et. al., 2001).The growth of methylotrophic microorganisms presents a challenge since they require molecular oxygen in the dissimilation of methanol, especially for the oxidation of methane to methanol. Yet, oxidation of methane under anaerobic conditions has been reported (Knittel and Boetius, 2009). Although it was a consortium of bacteria and *Archaea* that oxidized the methanol without molecular oxygen, the bacterial members of the consortia were also shown to oxidize methane in the absence of *Archaea*. Nakano and Zuber (1998) discussed the importance of nitrate reductase genes for the anaerobiosis of *B. subtilis*.

The examples mentioned above were grown on rich media using oxygenated compounds like nitrate as the terminal electron acceptor. However, the anaerobiosis of Wood-Ljungdahl pathway organisms are different from above mentioned microorganisms in that they fix CO_2 into biomass. The Wood-Ljungdahl pathway was briefly discussed in Chapter 1.

The *B. methanolicus* MGA3 genome was shown to harbor genes of the Wood-Ljungdahl pathway (4.13). Especially, the presence of carbon monoxide dehydrogenase (CODH) necessitated an investigation of possible anaerobiosis of *B. methanolicus*.

Materials and Methods

Growth of *B. methanolicus*: *B. methanolicus* MGA3 was grown overnight on mannitol in the presence of 25 mM sodium nitrate to acclimatize *B. methanolicus* to nitrate as an electron acceptor. Anaerobic serum bottles containing 50 ml of either $MTYM_{MeOH}$ or $MTYM_{MAN}$ media were prepared by flushing the head space with nitrogen gas for 5 minutes. The culture grown in the presence of 25 mM nitrate was inoculated (with varying initial concentrations) into the bottles using a syringe and then incubated at 50 °C. For comparison, bottles with aerobic head spaces were also inoculated with the *B. methanolicus* MGA3. For possible acclimatization, *B. methanolicus* MGA3 was grown with aerobic (air) headspace and allowed to deplete the oxygen in the headspace. The growth was monitored by measuring OD_{600} by taking samples from the bottle using syringes. The change in gas composition was measured via gas chromatography.

Results and conclusion

As seen from the Figure 7.1, death of *B. methanolicus* MGA3 starts immediately on inoculation of the seed culture into the anaerobic bottle. However the culture in the bottle with aerobic headspace has limited growth (OD_{600} of 1.336) until the oxygen was depleted. No further growth was observed after this point.

The head space gas compositions of aerobic bottles with either $MTYM_{MeOH}$(Figure 7.2) or $MTYM_{MAN}$(Figure 7.3) were monitored for any change over 72 hours after inoculation of *B. methanolicus* to OD_{600} of 1. In both bottles, the oxygen in the headspace was depleted in 6 hours and the CO_2 concentration peaked at the same time. The bottles were monitored for 72 h without any change in the composition of headspace gas, which indicates no growth of *B. methanolicus*. The bottles were incubated for an extended period of time (30 days). The culture during this period stayed dormant as demonstrated by the fact that one half a milliliter of culture from the bottle was able to start aerobic growth when transferred into 50 ml $MTYM_{MeOH}$ in shake flasks (data not shown). This might be the spores *B. methanolicus* formed during oxygen depletion since it is known to form spores. With the aim of acclimatizing *B. methanolicus* MGA to anaerobiosis, 5 ml of liquid from the anaerobic bottles were transferred to fresh anaerobic bottles every other day. Limited growth (OD_{600} of 0.350) was observed after several passages. The total DNA from this culture was isolated as described in Chapter 3. The primers designed to amplify the entire 16S rRNA of *B. methanolicus* MGA3 failed to amplify the target. Another primer set, which amplifies a shorter (~500 bp) fragment of 16S rRNA, amplified an expected size fragment but sequencing of the amplicon identified the microorganism as *Bacillus sonorensis*. The sequence is 99% identical to *B. sonorensis* while it matched 97% to *B. methanolicus* MGA3 16S rRNA. This is most probably a contamination introduced during passages. The trials of growing *B. methanolicus* MGA3 with different media (SOBsuc, 2XYT) and different gas composition in the head space (N_2, Ar, Syngas) were also unsuccessful.

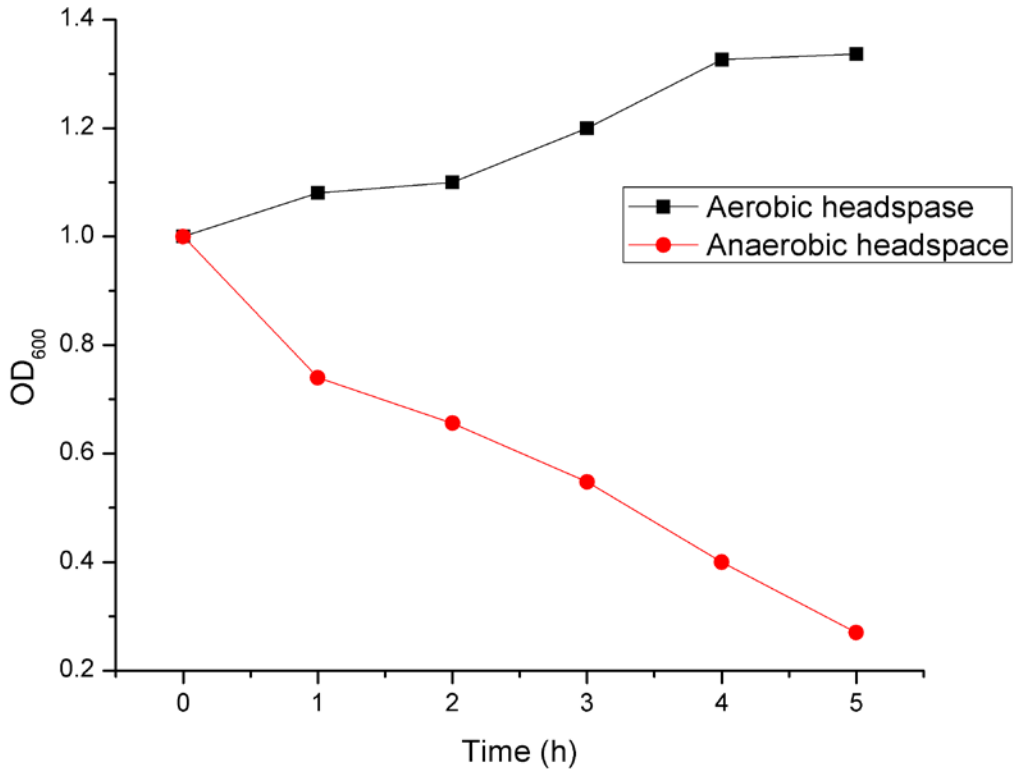


Figure 7.1. The growth of *B. methanolicus* MGA3 in serum bottles ($MTYM_{MAN}$) under aerobic (air) and anaerobic (N_2) head space. The bottles were inoculated to final OD_{600} of 1 and the growth was monitored for 5 hours. (n=1)

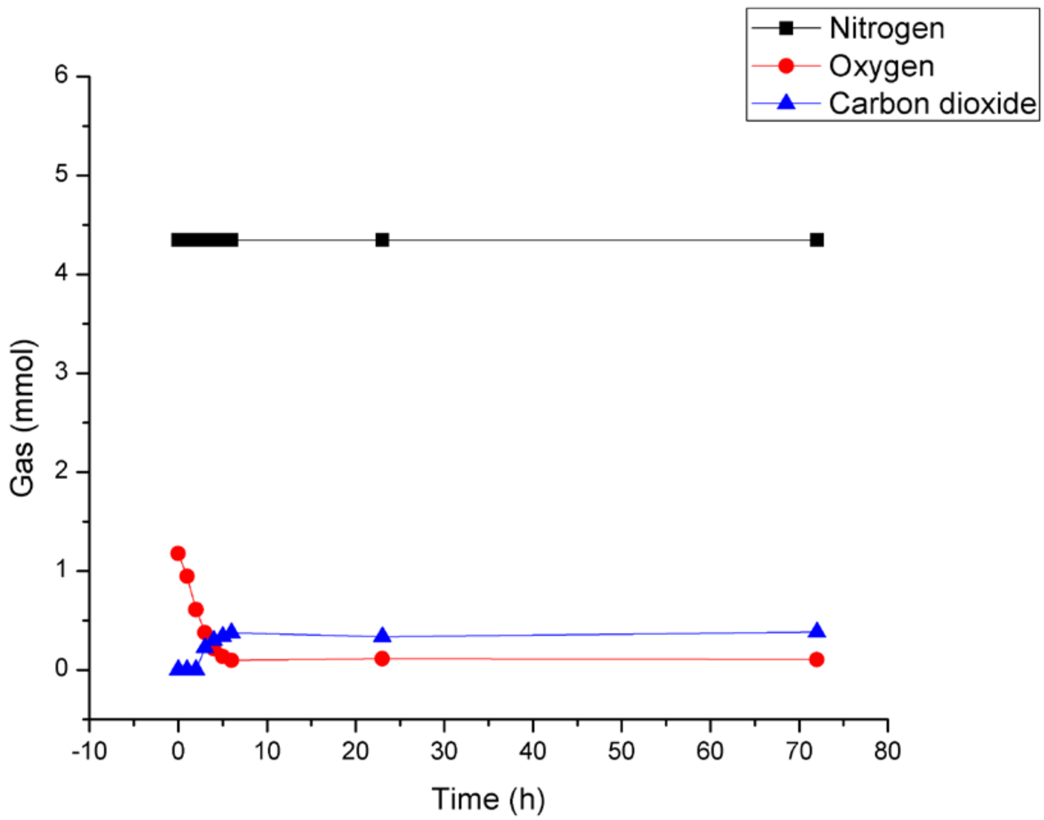


Figure 7.2. Analysis of gas composition of *B. methanolicus* MGA3 grown in $MTYM_{MeOH}$. *B. methanolicus* MGA3 was grown in $MTYM_{MeOH}$ in shake flasks aerobically as described in Chapter 2, in the presence of 25 mM sodium nitrate and inoculated to a final OD_{600} of 1.

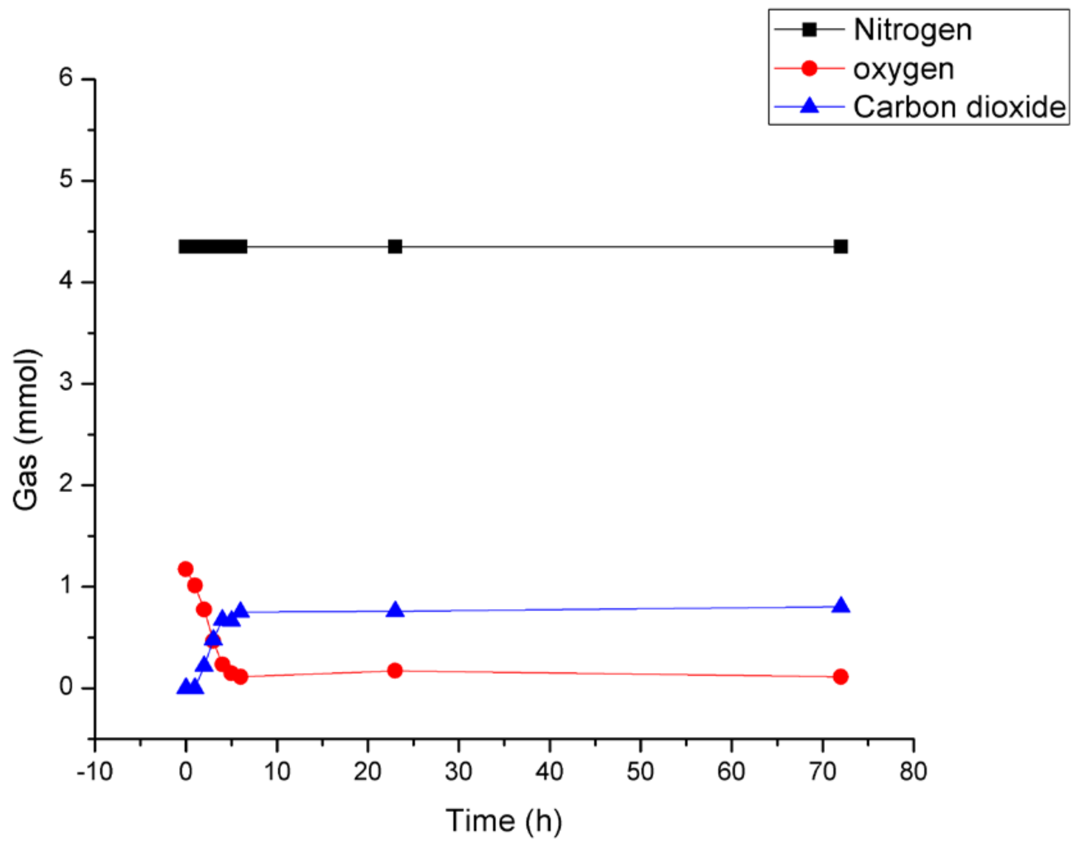


Figure 7.3. Analysis of gas composition of *B. methanolicus* MGA3 grown in $MTYM_{MAN}$. *B. methanolicus* MGA3 was grown in $MTYM_{MAN}$ in shake flasks aerobically as described in Chapter 2 in the presence of 25 mM sodium nitrate and inoculated into the bottle to final OD_{600} of 1.

CHAPTER 8

STUDY CONCLUSIONS

Bacillus methanolicus is a gram-positive, thermotolerant, restrictive methylotroph capable of using few substrates other than methanol like mannitol and glucose. The methylotrophy in *B. methanolicus* is known to be encoded by a plasmid. *B. methanolicus* employs the RuMP pathway for assimilation of carbon from formaldehyde. The formaldehyde is either assimilated into biomass via the RuMP pathway or dissimilated to CO_2. The rate of CO_2dissimilation increases when fed-batch cultures of *B. methanolicus* are subjected to a methanol or formaldehyde spike.

In this study we demonstrated that growth on methanol is challenging due to the possible accumulation of the metabolically important compound formaldehyde, despite the higher growth rate of *B. methanolicus*when grown on methanol than on mannitol or glucose. The growth rate on methanol is highest at methanol concentrations of 10 mM and 25 mM and then decreases as the methanol concentration is elevated (0.65±007 h^{-1} at 25 mM to 0.231±0.004 h^{-1} at 2 M). We have demonstrated that *B. methanolicus* growing on methanol can tolerate higher concentrations of methanol spike when compared to mannitol grown cells. However, it is interesting to note the growth inhibition when mannitol grown *B. methanolicus*cells were subjected to a methanol spike.*B. methanolicus* can utilize both substrates, and it can grow on 2 M methanol with a growth rate of 0.231±0.004 h^{-1}, yet 50 mM methanol was enough to perturb the growth of *B. methanolicus* when grown on mannitol. In this study we have also shown that methanol grown *B. methanolicus* is more susceptible to formaldehyde toxicity than mannitol grown cells probably due to the already high intracellular levels of formaldehyde in methanol grown cells. The detoxification pathways of formaldehyde assimilation and dissimilation, are saturated in methanol grown cells, and further addition of formaldehyde increases the level to toxic levels. The methanol metabolite formate was shown not to have any visible effect on the growth of either methanol or mannitol grown *B. methanolicus*.

B. methanolicus has a genome of ~3.3 Mb with the GC content of 38.35%. Its 2773 ORFs which are longer than 100 bp were annotated. Interestingly, duplicates of RuMP pathway genes were found on the chromosome. Also, putative genes of formaldehyde dissimilation were identified on the genome. It carries two additional *mdh* genes besides the gene on the pBM19; however, the plasmid-borne *mdh* is responsible for most ofthe methanol dehydrogenase activity as the RT-qPCR results demonstrate. No gene was annotated as a formaldehyde dehydrogenase while a number of putative aldehyde dehydrogenases were identified. The genome carries two putative formate dehydrogenases and the THF-linked pathway for the oxidation of formaldehyde. Surprisingly, it also harbors genes of the Wood-Ljungdahl pathway, including CO dehydrogenase and methylene-THF reductase.

In this study we demonstrate that the formaldehyde dissimilation metabolisms of *B. methanolicus* MGA3 and PB1 are different. The Southern Blots using formaldehyde and formate dehydrogenases of other bacilli demonstrated the difference between the metabolism of *B. methanolicus* MGA3 and PB1. The formaldehyde dissimilation pathway of strain PB1 is more efficient than that of MGA3, as it has higher formaldehyde and formate dehydrogenase activity in the crude cell extract.

We have shown the down regulation of methanol dissimilation pathway genes by methanol, formaldehyde and formate. In this regard, formate is more potent as a regulator despite the fact that it does not have an inhibitory effect on the growth of *B. methanolicus*. The expression of methanol dehydrogenase, formaldehyde and formate dehydrogenases, and THF-linked enzymes are all down regulated by methanol and its metabolites. Interestingly, the *mdh* in mannitol-grown *B. methanolicus* is also down regulated in response to methanol. But, the RuMP pathway genes, i.e *hps* and *pfk*, are upregulated in response to methanol and its metabolites in mannitol grown *B. methanolicus*. The expression level of *hps* is comparable to that of *mdh*, even though the later is encoded by a plasmid with 10-16 copies per cell. However, the *hps* is down regulated in methanol-grown *B. methanolicus* from its already high levels in response to formaldehyde and formate.

Methanol and formaldehyde metabolisms and their regulation were investigated in this study. This is the first study to report the effects of methanol, formaldehyde and formate when spiked in growing *B. methanolicus*cultures rather than comparison of methanol and mannitol grown cells.

REFERENCES

1. **Adeosun, EK, Smith, TJ, Hoberg, AM, Velarde, G, Ford, R, Dalton, H.** 2004. Formaldehyde dehydrogenase preparations from *Methylococcus capsulatus* (Bath) comprise methanol dehydrogenase and methylene tetrahydromethanopterin dehydrogenase. Microbiology. **150:**707-713.

2. **Ali, H, Murrell, JC.** 2009. Development and validation of promoter-probe vectors for the study of methane monooxygenase gene expression in *Methylococcus capsulatus* Bath. Microbiology. **155:**761-771. doi: 10.1099/mic.0.021816-0.

3. **Ali, H, Scanlan, J, Dumont, MG, Murrell, JC.** 2006. Duplication of the *mmoX* gene in *Methylosinus sporium*: cloning, sequencing and mutational analysis. Microbiology. **152:**2931-2942. doi: 10.1099/mic.0.29031-0.

4. **Allais, J, Louktibi, A.** 1983. Oxidation of methanol by the yeast *Pichia pastoris*. Purification and properties of the formate dehydrogenase. Agr. Biol. Chem. **47:**2547.

5. **Anesti, V, McDonald, IR, Ramaswamy, M, Wade, WG, Kelly, DP, Wood, AP.** 2005. Isolation and molecular detection of methylotrophic bacteria occurring in the human mouth. Environ. Microbiol. **7:**1227-1238. doi: 10.1111/j.1462-2920.2005.00805.x.

6. **Anthony, C.** 1982. The biochemistry of methylotrophs. Academic Press, London.

7. **Anthony, C.** 1978. The prediction of growth yields in methylotrophs. Journal of General Microbiology. **104:**91-104. doi: 10.1099/00221287-104-1-91.

8. **Anthony, C.** 1991. Assimilation of carbon by methylotrophs, p. 79-109. *In* I. Goldberg and J. S. Rokem (eds.), Biology of methylotrophs.

9. **Anthony, C, Williams, P.** 2003. The structure and mechanism of methanol dehydrogenase. Biochimica Et Biophysica Acta - Proteins & Proteomics. **1647:**18-23. doi: 10.1016/S1570-9639(03)00042-6.

10. **Antony, CP, Kumaresan, D, Ferrando, L, Boden, R, Moussard, H, Scavino, AF, Shouche, YS, Murrell, JC.** 2010. Active methylotrophs in the sediments of Lonar Lake, a saline and alkaline ecosystem formed by meteor impact. ISME Journal: Multidisciplinary Journal of Microbial Ecology. **4:**1470-1480. doi: 10.1038/ismej.2010.70.

11. **Arfman, N, de Vries, KJ, Moezelaar, HR, Attwood, MM, Robinson, GK, van Geel, M, Dijkhuizen, L.** 1992. Environmental regulation of alcohol metabolism in thermotolerant methylotrophic *Bacillus* strains. Arch. Microbiol. **157:**272-278.

12. **Arfman, N, Dijkhuizen, L, Kirchhof, G, Ludwig, W, Schleifer, KH, Bulygina, ES, Chumakov, KM, Govorukhina, NI, Trotsenko, YA, White, D.** 1992. *Bacillus methanolicus* sp. nov., a new species of thermotolerant, methanol-utilizing, endospore-forming bacteria. Int. J. Syst. Bacteriol. **42:**439-445.

13. **Arfman, N, Hektor, HJ, Bystrykh, LV, Govorukhina, NI, Dijkhuizen, L, Frank, J.** 1997. Properties of an NAD(H)-containing methanol dehydrogenase and its activator protein from *Bacillus methanolicus*. Eur. J. Biochem. **244:**426-433.

14. **Arfman, N, Van Beeumen, J, De Vries, GE, Harder, W, Dijkhuizen, L.** 1991. Purification and characterization of an activator protein for methanol dehydrogenase from thermotolerant *Bacillus* spp. Journal of Biological Chemistry. **266:**3955-3960.

15. **Atkinson, R.** 1990. Gas-phase tropospheric chemistry of organic compounds: a review. Atmospheric Environment.Part A.General Topics. **24:**1-41.

16. **Azra Moosvi, S, McDonald, IR, Pearce, DA, Kelly, DP, Wood, AP.** 2005. Molecular detection and isolation from Antarctica of methylotrophic bacteria able to grow with methylated sulfur compounds. Syst. Appl. Microbiol. **28:**541-554. doi: 10.1016/j.syapm.2005.03.002.

17. **Baskaya, FS, Zhao, X, Flickinger, MC, Wang, P.** 2010. Thermodynamic feasibility of enzymatic reduction of carbon dioxide to methanol. Appl. Biochem. Biotechnol. **162:**391-398. doi: 10.1007/s12010-009-8758-x.

18. **Binnewies, TT, Motro, Y, Hallin, PF, Lund, O, Dunn, D, La, T, Hampson, DJ, Bellgard, M, Wassenaar, TM, Ussery, DW.** 2006. Ten years of bacterial genome sequencing: comparative-genomics-based discoveries Funct. Integr. Genomics. **6:**165-185. doi: 10.1007/s10142-006-0027-2.

19. **Boden, R, Thomas, E, Savani, P, Kelly, DP, Wood, AP.** 2008. Novel methylotrophic bacteria isolated from the River Thames (London, UK). Environ. Microbiol. **10:**3225-3236. doi: 10.1111/j.1462-2920.2008.01711.x.

20. **Bolt, HM.** 1987. Experimental toxicology of formaldehyde. J. Cancer Res. Clin. Oncol. **113:**305-309.

21. **Boone, DR, Liu, Y, Zhao, Z, Balkwill, DL, Drake, GR, Stevens, TO, Aldrich, HC.** 1995. *Bacillus infernus* sp. nov., an Fe (III)-and Mn (IV)-reducing anaerobe from the deep terrestrial subsurface. Int. J. Syst. Bacteriol. **45:**441-448.

22. **Brautaset, T, Jakobsen, MOM, Flickinger, MC, Valla, S, Ellingsen, TE.** 2004. Plasmid-dependent methylotrophy in thermotolerant *Bacillus methanolicus*. J. Bacteriol. **186:**1229-1238.

23. **Brautaset, T, Jakobsen, OM, Degnes, KF, Netzer, R, Naerdal, I, Krog, A, Dillingham, R, Flickinger, MC, Ellingsen, TE.** 2010. *Bacillus methanolicus* pyruvate carboxylase and homoserine dehydrogenase I and II and their roles for L-lysine production from methanol at 50 ^{0}C. Appl. Microbiol. Biotechnol. **87:**951-964. doi: 10.1007/s00253-010-2559-6.

24. **Brautaset, T, Jakobsen, OM, Josefsen, KD, Flickinger, MC, Ellingsen, TE.** 2007. *Bacillus methanolicus*: a candidate for industrial production of amino acids from methanol at 50 degrees C. Appl. Microbiol. Biotechnol. **74:**22-34. doi: 10.1007/s00253-006-0757-z.

25. **Brautaset, T, Williams, MD, Dillingham, RD, Kaufmann, C, Bennaars, A, Crabbe, E, Flickinger, MC.** 2003. Role of the *Bacillus methanolicus* citrate synthase II gene, citY, in regulating the secretion of glutamate in L-lysine-secreting mutants. Appl. Environ. Microbiol. **69:**3986-3995.

26. **Brondino, CD, Passeggi, MC, Caldeira, J, Almendra, MJ, Feio, MJ, Moura, JJ, Moura, I.** 2004. Incorporation of either molybdenum or tungsten into formate dehydrogenase from *Desulfovibrio alaskensis* NCIMB 13491; EPR assignment of the proximal iron-sulfur cluster to the pterin cofactor in formate dehydrogenases from sulfate-reducing bacteria. J. Biol. Inorg. Chem. **9:**145-151. doi: 10.1007/s00775-003-0506-z.

27. **Chen, PT, Shaw, J, Chao, Y, David Ho, T, Yu, S.** 2010. Construction of chromosomally located T7 expression system for production of heterologous secreted proteins in *Bacillus subtilis*. J. Agric. Food Chem. **58:**5392-5399.

28. **Chinnawirotpisan, P, Theeragool, G, Limtong, S, Toyama, H, Adachi, OO, Matsushita, K.** 2003. Quinoprotein alcohol dehydrogenase is involved in catabolic acetate production, while NAD-dependent alcohol dehydrogenase in ethanol assimilation in *Acetobacter pasteurianus* SKU1108. J. Biosci. Bioengin. **96:**564-571. doi: 10.1016/S1389-1723(04)70150-4.

29. **Chistoserdov, AY, Chistoserdova, LV, McIntire, WS, Lidstrom, ME.** 1994. Genetic organization of the mau gene cluster in *Methylobacterium extorquens* AM1: complete nucleotide sequence and generation and characteristics of mau mutants. J. Bacteriol. **176:**4052-4065.

30. **Chistoserdova, L, Crowther, GJ, Vorholt, JA, Skovran, E, Portais, JC, Lidstrom, ME.** 2007. Identification of a fourth formate dehydrogenase in *Methylobacterium*

*extorquens*AM1 and confirmation of the essential role of formate oxidation in methylotrophy. J. Bacteriol. **189:**9076-9081. doi: 10.1128/JB.01229-07.

31. **Chistoserdova, L, Kalyuzhnaya, MG, Lidstrom, ME.** 2009. The expanding world of methylotrophic metabolism. Annu. Rev. Microbiol. **63:**477-499. doi: 10.1146/annurev.micro.091208.073600.

32. **Chistoserdova, L, Laukel, M, Portais, JC, Vorholt, JA, Lidstrom, ME.** 2004. Multiple formate dehydrogenase enzymes in the facultative methylotroph *Methylobacterium extorquens* AM1 are dispensable for growth on methanol. J. Bacteriol. **186:**22-28.

33. **Choi, YJ, Morel, L, Bourque, D, Mullick, A, Massie, B, Miguez, CB.** 2006. Bestowing inducibility on the cloned methanol dehydrogenase promoter (PmxaF) of *Methylobacterium extorquens* by applying regulatory elements of *Pseudomonas putida* F1. Appl. Environ. Microbiol. **72:**7723-7729. doi: 10.1128/AEM.02002-06.

34. **Chrastil, J, Wilson, JT.** 1975. A sensitive colorimetric method for formaldehyde. Anal. Biochem. **63:**202-207.

35. **Cordwell, S,J.** 1999. Microbial genomes and "missing" enzymes: redefining biochemical pathways. Arch. Microbiol. 269-279. doi: 10.1007/s002030050780.

36. **Crosby, RM, Richardson, KK, Craft, TR, Benforado, KB, Liber, HL, Skopek, TR.** 1988. Molecular analysis of formaldehyde-induced mutations in human lymphoblasts and *E. coli*. Environ. Mol. Mutagen. **12:**155-166.

37. **Cross, AB, Anthony, C.** 1980. The electron-transport chains of the obligate methylotroph *Methylophilus methylotrophus*. Biochem. J. **192:**429-439.

38. **Crowther, GJ, Kosály, G, Lidstrom, ME.** July 15, 2008. Formate as the main branch point for methylotrophic metabolism in *Methylobacterium extorquens* AM1. J. Bacteriol. **190:**5057-5062. doi: 10.1128/JB.00228-08.

39. **Cue, D, Lam, H, Dillingham, RL, Hanson, RS, Flickinger, MC.** 1997. Genetic manipulation of *Bacillus methanolicus*, a gram-positive, thermotolerant methylotroph. Appl. and Envir. Microbiol. **63:**1406-1420.

40. **Cue, D, Lam, H, Hanson, RS, Flickinger, MC.** 1996. Characterization of a restriction-modification system of the thermotolerant methylotroph *Bacillus methanolicus*.. Appl. Envir. Microbiol. **62:**1107-1111.

41. **de Vries, GE, Arfman, N, Terpstra, P, Dijkhuizen, L.** 1992. Cloning, expression, and sequence analysis of the *Bacillus methanolicus* C1 methanol dehydrogenase gene. J. Bacteriol. **174:**5346-5353.

42. **Degenkolbe, E.** 2010. Expression, Purification and Properties of Formaldehyde Dehydrogenase from different *Bacillus* strains. Master. University of Applied Sciences, Berlin.

43. **Dijkhuizen, L, Arfman, N, Attwood, MM, Brooke, AG, Harder, W, Watling, EM.** 1988. Isolation and initial characterization of thermotolerant methylotrophic *Bacillus* strains. FEMS Microbiol. Lett. **52:**209-214. doi: 10.1111/j.1574-6968.1988.tb02597.x.

44. **Dijkhuizen, L, Levering, P, de Vries, GE.** 1992. The physiology and biochemistry of aerobic methanol-utilizing gram-negative and gram-positive bacteria.*In* C. J. Murrell and H. Dalton (eds.), Methane and Methanol Utilizers.

45. **Doronina, NV, Darmaeva, TD, Trotsenko, YA.** 2003. *Methylophaga alcalica* sp. nov., a novel alkaliphilic and moderately halophilic, obligately methylotrophic bacterium from an East Mongolian saline soda lake. Int. J. Syst. and Evolutionary Microbiol. **53:**223-229. doi: 10.1099/ijs.0.02267-0.

46. **Duine, JA, Frank, J, Westerling, J.** 1978. Purification and properties of methanol dehydrogenase from *Hyphomicrobium* X. Biochim. Biophys. Acta. **524:**277-287.

47. **Enoch, HG, Lester, RL.** 1974. The role of a novel cytochrome b-containing nitrate reductase and quinone in the in vitro reconstruction of formate-nitrate reductase activity of *E. coli*. Biochem. Biophys. Res. Commun. **61:**1234-1241.

48. **Enoch, HG, Lester, RL.** 1975. The purification and properties of formate dehydrogenase and nitrate reductase from *Escherichia coli*. J. Biol. Chem. **250:**6693-6705.

49. **Erb, TJ, Berg, IA, Brecht, V, Muller, M, Fuchs, G, Alber, BE.** 2007. Synthesis of C5-dicarboxylic acids from C2-units involving crotonyl-CoA carboxylase/reductase: the ethylmalonyl-CoA pathway. Proc. Natl. Acad. Sci. U. S. A. **104:**10631-10636. doi: 10.1073/pnas.0702791104.

50. **Ettwig, KF, Shima, S, Van De Pas-Schoonen, KT, Kahnt, J, Medema, MH, Op Den Camp, HJM, Jetten, MSM, Strous, M.** 2008. Denitrifying bacteria anaerobically oxidize methane in the absence of Archaea. Environ. Microbiol. **10:**3164-3173. doi: 10.1111/j.1462-2920.2008.01724.x.

51. **Ferry, JG.** 1999. Enzymology of one-carbon metabolism in methanogenic pathways. FEMS Microbiol. Rev. **23:**13-38. doi: 10.1111/j.1574-6976.1999.tb00390.x.

52. **Fraser, CM, Eisen, JA, Nelson, KE, Paulsen, IT, Salzberg, SL.** 2002. The Value of Complete Microbial Genome Sequencing (You Get What You Pay For). Journal of Bacteriology. **184:**6403-6405. doi: 10.1128/JB.184.23.6403-6405.2002.

53. **Girio, FM, Roseiro, JC, Silva, AI.** 1998. The Effect of the simultaneous addition of molybdenum and tungsten to the culture medium on the formate dehydrogenase activity from *Methylobacterium sp*. RXM. Curr. Microbiol. **36:**337-340.

54. **Glaser, P, Danchin, A, Kunst, F, Zuber, P, Nakano, MM.** 1995. Identification and isolation of a gene required for nitrate assimilation and anaerobic growth of *Bacillus subtilis*. J. Bacteriol. **177:**1112-1115.

55. **Goenrich, M, Bartoschek, S, Hagemeier, CH, Griesinger, C, Vorholt, JA.** 2002. A glutathione-dependent formaldehyde-activating enzyme (Gfa) from *Paracoccus denitrificans* detected and purified via two-dimensional proton exchange NMR spectroscopy. J. Biol. Chem. **277:**3069-3072. doi: 10.1074/jbc.C100579200.

56. **Gutmacher, F.** 2009. Insigths into Carbon Flux in *Bacillus methanolicus*. M.S. Beuth University of Applied Sciences, Berlin, Germany.

57. **Haber, CL, Allen, LN, Zhao, S, Hanson, RS.** 1983. Methylotrophic Bacteria: biochemical diversity and genetics. Science. **221:**1147-1153. doi: 10.1126/science.221.4616.1147.

58. **Hagemeier, CH, Chistoserdova, L, Lidstrom, ME, Thauer, RK, Vorholt, JA.** 2000. Characterization of a second methylene tetrahydromethanopterin dehydrogenase from *Methylobacterium extorquens* AM1. Eur. J. Biochem. **267:**3762-3769. doi: 10.1046/j.1432-1327.2000.01413.x.

59. **Handelsman, J.** 2004. Metagenomics: Application of Genomics to Uncultured Microorganisms. Microbiology and Molecular Biology Reviews. **68:**669-685. doi: 10.1128/MMBR.68.4.669-685.2004.

60. **Hanson, RS.** 1992. Introduction, p. 1-22. *In* J. C. Murrell and H. Dalton (eds.), Methane and methanol utilizers. Plenum Press, New York.

61. **Hanson, RS, Dillingham, R, Olson, P, Lee, GH, Cue, D, Schendel, FJ, Bremmon, C, Flickinger, MC.** 1996. Production of L-lysine and some other amino acids by mutants of *Bacillus methanolicus*. I, p. 227-236. *In* M. E. Lidstrom and F. R. Tabita (eds.), Microbial growth on C1 compounds. Kluwer.

62. **Hanson, RS, Hanson, TE.** 1996. Methanotrophic bacteria. Microbiol. Rev. **60:**439-471.

63. **Harms, N, Ras, J, Reijnders, WN, van Spanning, RJ, Stouthamer, AH.** 1996. S-formylglutathione hydrolase of *Paracoccus denitrificans* is homologous to human esterase D: a universal pathway for formaldehyde detoxification. J. Bacteriol. **178:**6296-6299.

64. **Hazeu, W, Bruyn, JC, Bos, P.** 1972. Methanol assimilation by yeasts. Arch. Microbiol. **87:**185-188.

65. **Heck, HD, Casanova, M, Starr, TB.** 1990. Formaldehyde toxicity: new understanding. Crit. Rev. Toxicol. **20:**397-426. doi: 10.3109/10408449009029329.

66. **Heggeset, TMB, Krog, A, Balzer, S, Wentzel, A, Ellingsen, TE, Brautaset, T.** 2012. Genome Sequence of Thermotolerant *Bacillus methanolicus*: Features and Regulation Related to Methylotrophy and Production of L-Lysine and L-Glutamate from Methanol. Appl. Environ. Microbiol. **78(15):**5170.

67. **Hektor, HJ, Kloosterman, H, Dijkhuizen, L.** 2000. Nicotinoprotein methanol dehydrogenase enzymes in Gram-positive methylotrophic bacteria. J Molec Catal B. **8:**103-109. doi: 10.1016/S1381-1177(99)00073-9.

68. **Hendrickson, EL, Beck, DA, Wang, T, Lidstrom, ME, Hackett, M, Chistoserdova, L.** 2010. Expressed genome of *Methylobacillus flagellatus* as defined through comprehensive proteomics and new insights into methylotrophy. J. Bacteriol. **192:**4859-4867. doi: 10.1128/JB.00512-10.

69. **Intergovernmental Panel on Climate Change (IPCC) Climate change.** 1995. The science of climate: Parameterisation of change. Cambridge University Press. Cambridge, UK.*In* Anonymous . Cambridge University Press, Cambridge, UK.

70. **Jakobsen, OM, Benichou, A, Flickinger, MC, Valla, S, Ellingsen, TE, Brautaset, T.** 2006. Upregulated transcription of plasmid and chromosomal ribulose monophosphate pathway genes is critical for methanol assimilation rate and methanol tolerance in the methylotrophic bacterium *Bacillus methanolicus*. J. Bacteriol. **188:**3063-3072. doi: 10.1128/JB.188.8.3063-3072.2006.

71. **Jakobsen, OM, Brautaset, T, Degnes, KF, Heggeset, TM, Balzer, S, Flickinger, MC, Valla, S, Ellingsen, TE.** 2009. Overexpression of wild-type aspartokinase increases L-lysine production in the thermotolerant methylotrophic bacterium *Bacillus methanolicus*. Appl. Environ. Microbiol. **75:**652-661. doi: 10.1128/AEM.01176-08.

72. **Jormakka, M, Byrne, B, Iwata, S.** 2003. Formate dehydrogenase: a versatile enzyme in changing environments. Curr. Opin. Struct. Biol. **13:**418-423.

73. **Kalyuzhnaya, MG, Lapidus, A, Ivanova, N, Copeland, AC, McHardy, AC, Szeto, E, Salamov, A, Grigoriev, IV, Suciu, D, Levine, SR, Markowitz, VM, Rigoutsos, I, Tringe, SG, Bruce, DC, Richardson, PM, Lidstrom, ME, Chistoserdova, L.** 2008. High-resolution metagenomics targets specific functional types in complex microbial communities. Nat. Biotechnol. **26:**1029-1034. doi: 10.1038/nbt.1488.

74. **Kalyuzhnaya, MG, Lidstrom, ME.** 2003. QscR, a LysR-type transcriptional regulator and CbbR homolog, is involved in regulation of the serine cycle genes in *Methylobacterium extorquens* AM1. J. Bacteriol. **185:**1229-1235.

75. **Khangulov, SV, Gladyshev, VN, Dismukes, GC, Stadtman, TC.** 1998. Selenium-containing formate dehydrogenase H from *Escherichia coli*: a molybdopterin enzyme that catalyzes formate oxidation without oxygen transfer. Biochemistry (N. Y.). **37:**3518-3528.

76. **Kitchens, JF, Casner, RE, Edwards, GS, Harward, WE, Marci, BJ.** 1976. Investigation of selected potential environmental contaminants: Formaldehyde. US Environmental Protection Agency. Washington, DC. .

77. **Knittel, K, Boetius, A.** 2009. Anaerobic Oxidation of Methane: Progress with an unknown process. Annu. Rev. Microbiol. **63:**311-334. doi: 10.1146/annurev.micro.61.080706.093130.

78. **Komives, CF, Cheung, LY, Pluschkell, SB, Flickinger, MC.** 2005. Growth of *Bacillus methanolicus* in seawater-based media. J. Ind. Microbiol. Biotechnol. **32:**61-66. doi: 10.1007/s10295-004-0195-9.

79. **Krasner, S, McGuire, M, Jacangelo, J.** 1989. The occurrence of disinfection by-products in US drinking water Journal American Water Works Association. **81 (8):**41.

80. **Kuwahara, H.** 1980. Metabolism of lignin-related compounds by bacteria in lignin biodegradation, p. 148. *In* T. K. Kirk, T. Higuchi, and H. and Chang (eds.), Lignin Biodegradation: Microbiology, Chemistry and Potential Applicationsvol. 2. CRC Press, New York.

81. **Large, PJ, Peel, D, Quayle, JR.** 1961. Microbial growth on C1 compounds. II. Synthesis of cell constituents by methanol- and formate-grown *Pseudomonas* AM 1, and methanol-grown *Hyphomicrobium vulgare*. Biochem. J. **81:**470-480.

82. **Lee, JD, Komagata, K.** 1980. Taxonomic study of methanol-assimilating yeasts. J. Gen. Appl. Microbiol. **26:**133.

83. **Leopoldini, M, Russo, N, Toscano, M.** 2007. The preferred reaction path for the oxidation of methanol by PQQ-containing methanol dehydrogenase: addition-elimination

versus hydride-transfer mechanism. Chemistry. **13:**2109-2117. doi: 10.1002/chem.200601123.

84. **Lidstrom, ,M.E.** 2006. Aerobic methylotrophs.*In* M. Dworkin, S. Falkow, E. Rosenberg, K. Schleifer, and E. Stackebrandt (eds.), The prokaryotes: a handbook on the biology of bacteria. Springer, New York, New York.

85. **Lin, C, Liu, W, and Tan, C.** 2003. Removal of carbon dioxide by absorption in a rotating packed bed. Ind. Eng. Chem. Res. 2381. doi: 10.1021/ie020669+.

86. **Maniatis, T, Fritsch, EF, Sambrook, J.** 1982. Molecular cloning: a laboratory manual Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.

87. **Martin, H, Murrell, JC.** 1995. Methane monooxygenase mutants of *Methylosinus trichosporium* constructed by marker-exchange mutagenesis. FEMS Microbiol. Lett. **127:**243-248. doi: 10.1016/0378-1097(95)00068-G.

88. **Marx, CJ, Chistoserdova, L, Lidstrom, ME.** 2003. Formaldehyde-Detoxifying Role of the Tetrahydromethanopterin-Linked Pathway in *Methylobacterium extorquens* AM1. Journal of Bacteriology. **185:**7160-7168. doi: 10.1128/JB.185.23.7160-7168.2003.

89. **McGhee, JD, von Hippel, PH.** 1975. Formaldehyde as a probe of DNA structure. I. Reaction with exocyclic amino groups of DNA bases. Biochemistry. **14:**1281-1296.

90. **McGhee, JD, von Hippel, PH.** 1975. Formaldehyde as a probe of DNA structure. II. Reaction with endocyclic imino groups of DNA bases. Biochemistry. **14:**1297-1303.

91. **McGhee, JD, von Hippel, PH.** 1977. Formaldehyde as a probe of DNA structure. 3. Equilibrium denaturation of DNA and synthetic polynucleotides. Biochemistry. **16:**3267-3276.

92. **McGhee, JD, von Hippel, PH.** 1977. Formaldehyde as a probe of DNA structure. IV. Mechanism of the initial reaction of formaldehyde with DNA. Biochemistry. **16:**3276-3293.

93. **Merchant Research & Consulting ltd.** 2012. Formaldehyde: 2011 World Market outlook and Forecast upto 2017. .

94. **Mozzanega, P.** 2010. Characterisation of the enzymes involved in methanol dissimilation in *Bacillus methanolicus*: Assessment of enzyme-based biohydrogen production for a cell-free biofuel cell. University of Bath.

95. **Muir, PS.** 1991. Fogwater chemistry in a wood-burning community, Western Oregon. J. Air Waste Manage. Assoc. **41:**32-38.

96. **Nakano, M, Zuber, P.** 1998. Anaerobic growth of a "strict aerobe" (*Bacillus subtilis*). Annu. Rev. Microbiol. **52:**165-190. doi: 10.1146/annurev.micro.52.1.165.

97. **National Research Council (U.S.). Committee on Toxicology.** 1980. Formaldehyde: an assesment of its health effects/ prepared for the Consumer Product Safety Commission by the committe on the Toxicology, Board on Toxicology and Environmental Health Hazards, Assembly of Life Sciences, National Research Council. Natinal Academy of Science. .

98. **National Toxicology Program.** 2010. Final Report on Carcinogens Background Document for Formaldehyde. Rep. Carcinog. Backgr Doc. **(10-5981):**1-512.

99. **Nercessian, O, Kalyuzhnaya, MG, Joye, SB, Lidstrom, ME, Chistoserdova, L.** 2005. Analysis of *fae* and *fhcD* Genes in Mono Lake, California. Appl. Envir. Microbiol. **71:**8949-8953. doi: 10.1128/AEM.71.12.8949-8953.2005.

100. **Neufeld, JD, Chen, Y, Dumont, MG, Murrell, JC.** 2008. Marine methylotrophs revealed by stable-isotope probing, multiple displacement amplification and metagenomics. Environ. Microbiol. **10:**1526-1535. doi: 10.1111/j.1462-2920.2008.01568.x.

101. **Newton, GL, Rawat, M, La Clair, JJ, Jothivasan, VK, Budiarto, T, Hamilton, CJ, Claiborne, A, Helmann, JD, Fahey, RC.** 2009. Bacillithiol is an antioxidant thiol produced in Bacilli. Nature Chemical Biology. **5:**625-627.

102. **Nguyen, TT, Eiamphungporn, W, Mader, U, Liebeke, M, Lalk, M, Hecker, M, Helmann, JD, Antelmann, H.** 2009. Genome-wide responses to carbonyl electrophiles in *Bacillus subtilis*: control of the thiol-dependent formaldehyde dehydrogenase AdhA and cysteine proteinase YraA by the MerR-family regulator YraB (AdhR). Mol. Microbiol. **71:**876-894. doi: 10.1111/j.1365-2958.2008.06568.x.

103. **Nilasari, D, Dover, N, Rech, S, Komives, C.** 2012. Expression of recombinant green fluorescent protein in *Bacillus methanolicus*. Biotechnol. Prog. . doi: 10.1002/btpr.1522.

104. **Oakley, CJ, Colin Murrell, J.** 1988. *nifH* genes in the obligate methane oxidizing bacteria. FEMS Microbiol. Lett. **49:**53-57. doi: 10.1111/j.1574-6968.1988.tb02681.x.

105. **Oakley, CJ, Colin Murrell, J.** 1991. Cloning of nitrogenase structural genes from the obligate methanotroph *Methylococcus capsulatus* (Bath). FEMS Microbiol. Lett. **78:**121-125.

106. **O'Brien, PJ, Siraki, AG, Shangari, N.** 2005. Aldehyde sources, metabolism, molecular toxicity mechanisms, and possible effects on human health. Crit. Rev. Toxicol. **35:**609-662.

107. **O'Connor, ML, Quayle, JR.** 1980. Pentose phosphate-dependent fixation of formaldehyde by methanol-grown *Hansenula polymorpha* and *Candida boidinii*. Microbiology. **120:**219. doi: 10.1099/00221287-120-1-219.

108. **Oelgeschlager, E, Rother, M.** 2008. Carbon monoxide-dependent energy metabolism in anaerobic bacteria and archaea. Arch. Microbiol. 257-269. doi: 10.1007/s00203-008-0382-6.

109. **Ogata, K, Nishikawa, H, Ohsugi, M.** 1969. A yeast capable of utilizing methanol. Agr. Biol. Chem. **33:**1519.

110. **Oki, T, Kouno, K, Kitai, A.** 1972. New yeasts capable of assimilating methanol J. Gen. Appl. Microbiol. **18:**295.

111. **Peyraud, R, Kiefer, P, Christen, P, Massou, S, Portais, JC, Vorholt, JA.** 2009. Demonstration of the ethylmalonyl-CoA pathway by using ^{13}C metabolomics. Proc. Natl. Acad. Sci. U. S. A. **106:**4846-4851. doi: 10.1073/pnas.0810932106.

112. **Pluschkell, SB.** 1998. Characterization and mathematical modeling of growth and glutamic acid production by *Bacillus methanolicus* MGA3. A thesis submitted to University of Minnesota. University of Minnesota.

113. **Pluschkell, SB, Flickinger, MC.** 2002. Dissimilation of ^{13}C methanol by continuous cultures of *Bacillus methanolicus* MGA3 at 50 ^{o}C studied by ^{13}C NMR and isotope-ratio mass spectrometry. Microbiology. **148:**3223-3233.

114. **Qin, J, Li, R, Raes, J, Arumugam, M, Burgdorf, KS, Manichanh, C, Nielsen, T, Pons, N, Levenez, F, Yamada, T, Mende, DR, Li, J, Xu, J, Li, S, Li, D, Cao, J, Wang, B, Liang, H, Zheng, H, Xie, Y, Tap, J, Lepage, P, Bertalan, M, Batto, J, Hansen, T, Le Paslier, D, Linneberg, A, Nielsen, HB, Pelletier, E, Renault, P, Sicheritz-Ponten, T, Turner, K, Zhu, H, Yu, C, Li, S, Jian, M, Zhou, Y, Li, Y, Zhang, X, Li, S, Qin, N, Yang, H, Wang, J, Brunak, S, Dore, J, Guarner, F, Kristiansen, K, Pedersen, O, Parkhill, J, Weissenbach, J, Bork, P, Ehrlich, SD, Wang, J.** 2010. A human gut microbial gene catalogue established by metagenomic sequencing. Nature. **464:**59-65.

115. **Radajewski, S, Webster, G, Reay, DS, Morris, SA, Ineson, P, Nedwell, DB, Prosser, JI, Murrell, JC.** 2002. Identification of active methylotroph populations in an acidic forest soil by stable-isotope probing. Microbiology. **148:**2331-2342.

116. **Raghoebarsing, AA, Pol, A, van de Pas-Schoonen, KT, Smolders, AJ, Ettwig, KF, Rijpstra, WI, Schouten, S, Damste, JS, Op den Camp, HJ, Jetten, MS, Strous, M.** 2006. A microbial consortium couples anaerobic methane oxidation to denitrification. Nature. **440:**918-921. doi: 10.1038/nature04617.

117. **Ragsdale, S.** 2004. Life with Carbon Monoxide. Crit. Rev. Biochem. Mol. Biol. **39:**165-195. doi: 10.1080/10409230490496577.

118. **Ragsdale, SW.** 2008. Enzymology of the Wood-Ljungdahl pathway of acetogenesis. Ann. N. Y. Acad. Sci. **1125:**129-136. doi: 10.1196/annals.1419.015.

119. **Ras, J, Van Ophem, PW, Reijnders, WN, Van Spanning, RJ, Duine, JA, Stouthamer, AH, Harms, N.** 1995. Isolation, sequencing, and mutagenesis of the gene encoding NAD- and glutathione-dependent formaldehyde dehydrogenase (GD-FALDH) from *Paracoccus denitrificans*, in which GD-FALDH is essential for methylotrophic growth. J. Bacteriol. **177:**247-251.

120. **Reuss, G, Disteldorf, W, Gamer, AO, Hilt, A.** 2005. Formaldehyde*In Ullmann's Encyclopedia of Industrial Chemistry*. Wiley-VCH Verlag GmbH & Co. KGaA.

121. **Riedel, K, Weller, R, Schrems, O.** 1999. Variability of formaldehyde in the Antarctic troposphere. Phys. Chem. Chem. Phys. **1:**5523-5527.

122. **Roger, A.** 1990. Gas-phase tropospheric chemistry of organic compounds: A review. Atmos. Environ. **24:**1-41. doi: 10.1016/0960-1686(90)90438-S.

123. **Sawers, G.** 1994. The hydrogenases and formate dehydrogenases of *Escherichia coli*. Antonie Van Leeuwenhoek. **66:**57-88.

124. **Schendel, FJ, Dillingham, R, Hanson, RS, Sano, K, Matsui, K.** 2000. Production of glutamate using wild-type *Bacillus methanolicus*. US patent 6083728.

125. **Schendel, FJ, Bremmon, CE, Flickinger, MC, Guettler, M, Hanson, RS.** 1990. L-lysine production at 50 ^{o}C by mutants of a newly isolated and characterized methylotrophic *Bacillus* sp. Appl. and Environ. Microbiol. **56:**963-970.

126. **Schink, B, Zeikus, J.** 1980. Microbial methanol formation: A major end product of pectin metabolism. Curr. Microbiol. 387-389. doi: 10.1007/BF02605383.

127. **Schrader, J, Schilling, M, Holtmann, D, Sell, D, Filho, MV, Marx, A, Vorholt, JA.** 2009. Methanol-based industrial biotechnology: current status and future perspectives of methylotrophic bacteria. Trends Biotechnol. **27:**107-115. doi: 10.1016/j.tibtech.2008.10.009.

128. **Schütte, H, Flossdorf, J, Sahm, H, Kula, M.** 1976. Purification and properties of formaldehyde dehydrogenase and formate dehydrogenase from *Candida boidinii*. Eur. J. Biochem. **62:**151-160.

129. **Seravalli, J, Ragsdale, SW.** 2000. Channeling of carbon monoxide during anaerobic carbon dioxide fixation. Biochemistry. **39 (6):**1274. doi: 10.1021/bi991812e.

130. **Simon, R, Priefer, U, Pühler, A.** 1983. A broad host range mobilization system for in vivo genetic engineering: transposon mutagenesis in gram negative bacteria. Bio/technology. **1:**784-791. doi: 10.1038/nbt1183-784.

131. **Smith, MJ.** 2007. Thesis. The Formaldehyde dissimilatory of *Bacilli*;Characterization of Single Carbon Dehydrogenases. University of Minnesota.

132. **Smith, TJ, Slade, SE, Burton, NP, Murrell, JC, Dalton, H.** 2002. Improved system for protein engineering of the hydroxylase component of soluble methane monooxygenase. Appl. Environ. Microbiol. **68:**5265-5273.

133. **Solomon, S, Plattner, G, Knutti, R, Friedlingstein, P.** 2009. Irreversible climate change due to carbon dioxide emissions. Proceedings of the National Academy of Sciences. . doi: 10.1073/pnas.0812721106.

134. **Stolyar, SM, Romanovskaya, VA, Malashenko, YR.** 1995. Search for systems of genetic exchange in methane-oxidizing bacteria. Microbiologya. **64 (5):**686-691.

135. **Sy, A, Giraud, E, Jourand, P, Garcia, N, Willems, A, de Lajudie, P, Prin, Y, Neyra, M, Gillis, M, Boivin-Masson, C, Dreyfus, B.** 2001. Methylotrophic *Methylobacterium* bacteria nodulate and fix nitrogen in symbiosis with legumes. J. Bacteriol. **183:**214-220. doi: 10.1128/JB.183.1.214-220.2001.

136. **Tettelin, H, Feldblyum, T.** 2009. Bacterial genome sequencing Methods Mol. Biol. **551:**231-247. doi: 10.1007/978-1-60327-999-4_18.

137. **Thompson, JD, Higgins, DG, Gibson, TJ.** 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. **22:**4673-4680.

138. **Tirado-Acevedo, O.** 2010. Production of bioethanol from synthesis gas using *Clostridium ljungdahlii* as a microbial catalyst. North Carolina State University.

139. **Tishkov, VI, Popov, VO.** 2006. Protein engineering of formate dehydrogenase. Biomol. Eng. **23:**89-110. doi: 10.1016/j.bioeng.2006.02.003.

140. **Topp, E, Knowles, R.** 1984. Effects of nitrapyrin [2-chloro-6-(trichloromethyl) pyridine] on the obligate methanotroph *Methylosinus trichosporium* OB3b. Appl. Environ. Microbiol. **47:**258-262.

141. **Trézl, L, Csiba, A, Juhász, S, Szentgyörgyi, M, Lombai, G, Hullán, L, Juhász, A.** 1997. Endogenous formaldehyde level of foods and its biological significance Zeitschrift f?r Lebensmitteluntersuchung Und -Forschung A. **205:**300-304. doi: 10.1007/s002170050169.

142. **Trotsenko, YA, Bystrykh, LV, Ubiyvovk, VM.** 1984. Regulatory aspects of methanol metabolism in yeast. In R. L. Crawford and R. S Hanson (ed.), Microbial growth on C1 compounds. American Society for Microbiology, Washington, DC. 118-122.

143. **Trotsenko, ,Yu., Doronina, ,N., Khmelenina, ,V.** 2005. Biotechnological potential of aerobic methylotrophic bacteria: A review of current state and future prospects. Appl. Biochem. Microbiol. 433-441. doi: 10.1007/s10438-005-0078-5.

144. **Tyson, GW, Chapman, J, Hugenholtz, P, Allen, EE, Ram, RJ, Richardson, PM, Solovyev, VV, Rubin, EM, Rokhsar, DS, Banfield, JF.** 2004. Community structure and metabolism through reconstruction of microbial genomes from the environment. Nature. **428:**37-43.

145. **Unden, G, Bongaerts, J.** 1997. Alternative respiratory pathways of *Escherichia coli*: energetics and transcriptional regulation in response to electron acceptors. Biochim. Biophys. Acta. **1320:**217-234.

146. **Van Dien, SJ, Lidstrom, ME.** 2002. Stoichiometric model for evaluating the metabolic capabilities of the facultative methylotroph *Methylobacterium extorquens* AM1, with application to reconstruction of C3 and C4 metabolism. Biotechnol. Bioeng. **78:**296-312. doi: 10.1002/bit.10200.

147. **van Ophem, PW, Van Beeumen, J, Duine, JA.** 1992. NAD-linked, factor-dependent formaldehyde dehydrogenase or trimeric, zinc-containing, long-chain alcohol dehydrogenase from *Amycolatopsis methanolica*. Eur. J. Biochem. **206:**511-518.

148. **Venter, JC, Remington, K, Heidelberg, JF, Halpern, AL, Rusch, D, Eisen, JA, Wu, D, Paulsen, I, Nelson, KE, Nelson, W, Fouts, DE, Levy, S, Knap, AH, Lomas, MW, Nealson, K, White, O, Peterson, J, Hoffman, J, Parsons, R, Baden-Tillson, H, Pfannkoch, C, Rogers, Y, Smith, HO.** 2004. Environmental genome shotgun sequencing of the Sargasso Sea. Science. **304:**66-74. doi: 10.1126/science.1093857.

149. **Vitousek, PM, Mooney, HA, Lubchenco, J, Melillo, JM.** 1997. Human domination of earth's ecosystems. Science. **277:**494-499. doi: 10.1126/science.277.5325.494.

150. **Vonck, J, Arfman, N, De Vries, GE, Van Beeumen, J, Van Bruggen, EF, Dijkhuizen, L.** 1991. Electron microscopic analysis and biochemical characterization of a novel methanol dehydrogenase from the thermotolerant *Bacillus sp.* C1. J. Biol. Chem. **266:**3949-3954.

151. **Vorholt, JA, Chistoserdova, L, Stolyar, SM, Thauer, RK, Lidstrom, ME.** 1999. Distribution of tetrahydromethanopterin-dependent enzymes in methylotrophic bacteria and phylogeny of methenyl tetrahydromethanopterin cyclohydrolases. J. Bacteriol. **181:**5750-5757.

152. **Vorholt, JA, Marx, CJ, Lidstrom, ME, Thauer, RK.** 2000. Novel formaldehyde-activating enzyme in *Methylobacterium extorquens* AM1 required for growth on methanol. J. Bacteriol. **182:**6645-6650.

153. **Vuilleumier, S, Chistoserdova, L, Lee, MC, Bringel, F, Lajus, A, Zhou, Y, Gourion, B, Barbe, V, Chang, J, Cruveiller, S, Dossat, C, Gillett, W, Gruffaz, C, Haugen, E, Hourcade, E, Levy, R, Mangenot, S, Muller, E, Nadalig, T, Pagni, M, Penny, C, Peyraud, R, Robinson, DG, Roche, D, Rouy, Z, Saenampechek, C, Salvignol, G, Vallenet, D, Wu, Z, Marx, CJ, Vorholt, JA, Olson, MV, Kaul, R, Weissenbach, J, Medigue, C, Lidstrom, ME.** 2009. Methylobacterium genome sequences: a reference blueprint to investigate microbial metabolism of C1 compounds from natural and industrial sources. PLoS One. **4:**e5584. doi: 10.1371/journal.pone.0005584.

154. **Wadzinski, AM, Ribbons, DW.** 1975. Oxidation of C1 compounds by particulate fractions from *Methylococcus capsulatus*: properties of methanol oxidase and methanol dehydrogenase. Journal of Bacteriology. **122:**1364-1374.

155. **Walker, JF.** 1975. Formaldehyde. Huntington, New York.

156. **Ward, N, Larsen, O, Sakwa, J, Bruseth, L, Khouri, H, Durkin, AS, Dimitrov, G, Jiang, L, Scanlan, D, Kang, KH, Lewis, M, Nelson, KE, Methe, B, Wu, M, Heidelberg, JF, Paulsen, IT, Fouts, D, Ravel, J, Tettelin, H, Ren, Q, Read, T, DeBoy, RT, Seshadri, R, Salzberg, SL, Jensen, HB, Birkeland, NK, Nelson, WC, Dodson, RJ, Grindhaug, SH, Holt, I, Eidhammer, I, Jonasen, I, Vanaken, S, Utterback, T, Feldblyum, TV, Fraser, CM, Lillehaug, JR, Eisen, JA.** 2004. Genomic insights into methanotrophy: the complete genome sequence of *Methylococcus capsulatus* (Bath). PLoS Biol. **2:**e303. doi: 10.1371/journal.pbio.0020303.

157. **Warner, PJ, Higgins, IJ, Drozd, JW.** 1980. Conjugative transfer of antibiotic resistance to methylotrophic bacteria. FEMS Microbiol. Lett. **7:**181-185. doi: 10.1111/j.1574-6941.1980.tb01621.x.

158. **West, CA, Salmond, GP, Dalton, H, Murrell, JC.** 1992. Functional expression in *Escherichia coli* of proteins B and C from soluble methane monooxygenase of *Methylococcus capsulatus* (Bath). J. Gen. Microbiol. **138:**1301-1307.

159. **World Health Organization.** 2005. Formaldehyde in Drinkingwater, Background Document for Development of WHO Guidelines for Drinking-Water Quality. .

160. **Xu, HH, Viebahn, M, Hanson, RS.** 1993. Identification of methanol-regulated promoter sequences from the facultative methylotrophic bacterium *Methylobacterium organophilum* XX. J. Gen. Microbiol. **139:**743-752. doi: 10.1099/00221287-139-4-743.

161. **Zahn, JA, Bergmann, DJ, Boyd, JM, Kunz, RC, DiSpirito, AA.** 2001. Membrane-associated quinoprotein formaldehyde dehydrogenase from *Methylococcus capsulatus* Bath. J. Bacteriol. **183:**6832-6840. doi: 10.1128/JB.183.23.6832-6840.2001.

162. **Zhang, X, Reddy, SY, Bruice, TC.** 2007. Mechanism of methanol oxidation by quinoprotein methanol dehydrogenase. Proc. Natl. Acad. Sci. U. S. A. **104:**pp. 745-749.

APPENDICES

Appendix A

Sequence of pBM20

AACGTCAAAAACCCTTATTTTTACTGGTTTTTTAGACGTCGAAAACAATAAAATA
TTTAAATATTTAAAATGTATTATTGTAAAACGACAAAACAAAGATGTATAAAAA
ATTTTTTTCATGTTATTTACTCAACTTTCGCAGAGTAATTTCTATATGCAACTCTT
GATACAACTGCTTTGCTTCGTCTTTCATTATTTCTCTTGTTTATAATGGACCCATA
AAACTAGACACGAAATGAGGAGATGTTAAAGTAGGTTTATGAAAAGAAAACGA
TATACACCAGAGTTTAAATCCCAGATCGTGCTAGAGGTCCTAAAAGAAGAAAAA
ACTATGAGTGAAATTGCTTCCTCTCATGGAATCCATGTGAATCAGATTCGGCAAT
GGAGGAATGCGTTTCTTGAACAAATGCCCCAGTTGTTTGCGAAAGAAAATAAGA
AAGTAGATCAAATGAAGGCTGACTATGAACAACAAATTGAAAACCTTTATGCCG
AAGTCGGTCGTTTAACCACGCAATTGTCGTGGCTTAAAAAAAAATCTGGCCTTAA
GGAGTAGGGCAGAACGTATTGACATGATTGATTGGGACGATGCAGAACTCCCTA
TTGCCACCCAGGCTGAATTGCTCGGACTCAATCGGTCCAGCTTGTATTATAAGCC
TGTGGAGCCTTCTCCTGAAGAAGTGTTCATCAAGCACCGAATTGATGAAATTTAC
ACCAAATATCCGTTTTTTGGCTCCCGGCGGATCACGGCCGTTCTGAATGAAAACG
GGTTACACATTAACCGAAAGGCCGTTCAGCGCCATATGAGGGAGATGGGGATTG
CCGGGATTTCACCTGGCCCCAACCTTAGTAAGCGCAACCTGGAACACCGGATTT
ATCCCTATCTTCTTAGAAATCTCGACATCACTCACCCCAACCAGGTATGGGGGAT
TGATATCACTTACATCCGGCTTCAACGCGGTTGGATGTATCTCGTCGCCATTATT
GACTGGTATTCCCGTTACGTCGTTAGTTGGGAGCTGGATCAGACGCTGGAAATCG
ATTTTGTCCTTGAAGCCGTCAAACGGGCGCTGTCGCACGCTCAACCTGAGATTAT
AAACAGTGATCAGGGTAGCCACTTCACCAGCCCGAAATATACCGAGCTGGTACT
GGATAAGAAAGTTCAAATCAGTATGGACGGCAAAGGTCGGGCTCTTGACAATAT
CATTACCGAGCGATTGTGGCGGACGATTAAATATGAAGAGGTCTACTTGAAGGA
ATACGAGAGTCCAAGAGAAGCAAGAAAAGAGATCAATAACTATCTCCATTTCTA
TAACCATGAACGCCCTCATCAATCATTAGGCTATAAACCACCTGCCTCTGTTTAT
GGCCTGTGAGTGATTCATCCTTTATAGAGGATCCTTCTCACGGGAGAACCTGTCA
AGGATTCGGCAGAGCCCTCGCTTCGCTCGCCCTTGACAGAACCTCCCCTGAGAA
GGAAAAATAAAAGCAGGATGAATCACTCCAAAAAAATGTGTACTCCCACTTTAT
TGGCCTCAATGAGCACCCCAAGCTCTAGAATGCAGATATCCTTTTTGGTTTCTGC
CCCACCTTAAAAAAATAAAATTAATGTCTTGATGATGGGGTCCACCATAGTTTTT
TATTGATTTTATTTTTTTAAAAAGAAAAAGACACCTATCTTTTGGTAAAATAATG
GTGACTAAACCAAAATACCAAAGAAGGTGTCACCATTATGATAACGAAAAAAGA
TACTTTTGTACAACTACCAAAAGAACTTGAACGAGGATTTTCTGAATTAAATATA
GGCAAACATTTGAGAAAAGCAAGAATTACTAAGGGATTCGGTTATTCATGCTTA
TCTATTTTCCGTCTTATCTTTTTGCTTGTTTTCCAATATAAGAACTGGTTTCAAGC
ACTTCAAAGTAAGAAAGCCGTGGATTTGCCTAAAAAAGATACCATTTACCGATTT
TTAAATTCATCTACTTATGCTTGGCGATCGTTTTTACTTTCTTTGTGTCACGACCTT

ATTATGCGTATAAAAAAGCTTACTTCTAACAAGAGAGTGAAGGTGTTTATTGTCG
ATGATTCACTCTTTTCACGGAATCGAAGCAAGTCTGTTGAACTTCTTTCTCGGGT
ATTTGGTCATACAGAGAAACGATTCTATTAAGTTCAGCGAAAAAGGAAAATCAA
CTTAACGGGATTAATGAATCCATTGATAAACGCACTTCCGGTTATAAACGTCGTC
TTGAAGCACAACTAGATAAGCCTAGCGCTATCTCAGCTTTATTAGATCATGCCTT
AAATGCTGGGATTACAGCTGATTATGTGTTAATGGATGCTTGGTTTACGCATGAG
CCTTTGATTGAAAAAATCACGGATAAAGGTTTATTTGTCATTGGTATGGTCAAGC
AGCTAAAACAACGATATTTCATAAATGGTAAAGCCCATACTCTGGATCAACTCTT
TAAAAAAGCAAAGCACACAATGGAAAAGAAAGAGCTACTCGGTTCGATCCATGT
TCATCTTGCGAAAAGACTGCAAGTTGTGTTTTGCAATTTAGAAATCCAGCGTTTT
GCAAGATTTTTTCCAGCACCTAACTGGAAATTCTGTAAAGGCATTTAAGAGTAAA
TGGGCTTGACACGGAGCCTCTGCGGACATGATTCTTTTGTCAAGATCCAGGCATA
ACAATGTTTGGCAAACTTGAACAAGGTTATCATGTCCGCTACTCCATGTAAAGCC
CACAAAATAAACTTAAAAAGTGCAATTGTTTTTTCCAGTTGGGTGCAGGATTTTT
ATTTTGCAAAATGCTCCATTTCTATTTTGCAATAAACACAAGTCAAGATCGTGTT
CGTGCGTAACCGCAATAAAAAGAGTGAATGGTTAGCTATTTTGAGTACAGATAC
CACACTCTCAGATGAAGAAATTATGCGTATTTACGGGATGCGCTGGGACATCGA
AACCTTTTTTAAATGTACAAAATCATTGCTTAACCTAGCAAAAGAATTTCAAGGT
CGTTCGTATAATCTGTTGATTAGCCATACAACAATTGTGTTTACTCGCTTTATACT
CTTAGAATGGGAACGTAGACAAAAGAACGATCCTAAGACAATTGGAAACCTCTT
TTTTCTTTTATGCGAAGACATAAAAGACATAGATCTAGAAAGTACGTTGAATCAA
TTATTAGTGACTTTCCAAGCACTAGCAGAAGCTAATGTTTGTTTGAACATGAAAT
TATTCAAAAATAAAGTACAAGAATGGATTGCGTCTTTACCTAATTACATCAAGGG
ATGCCTGACAATATCCGTGTGCGAAAGTTGAGTTAAAATAAAAAGGAGTTGCAA
AAAACGATCCCTCTTAACAAGTTTTAGTGTTGATAAATGAAAAAAGAAAGTGAT
TGGTGAAATAGTTTAATAGTGGCTAGTTGTATAATCAAATGCAACAAATCTTAAA
ACGTGAAAAACCATAGATAAACCGTGTATGATTAAGTCGACCAAAACAAAAATC
GATACGGAGGTTTTATCTATGGTTCAAACACAGTCTACCACAAAAAAACGCCTAT
TTAAACATCTTACACCCTTTGATCGCGGAAAAATAGCCGCATTACGTGCTGAGGG
AAAATCTCTTCAGGCTATCGCTGATGCGGTGGGCTGCCATAAAAGCACCATTTCT
CGAGAGTTAAAACGAGGCACCGTGACTCAGATGAAGACAGGTGGAAAGCTATTT
GAAGCTTATTTTCCGGATACGGGCCAACTGGTATACGAAAAGAACCGGAAAGCC
TGCAGTGCCAAATTTAAGTTGGATGACGCGATGGAATTCATCCAGTTTGCAGAG
GCCAAAATCCTAAGGGATGGCTGGTCACCTGATGCAGTCTGTGGTTATACCAAG
AGCCACGGTCTATTTGAAGGTGCCCTTATTTCTACCAAGACACTCTATAACTATA
TTGACATAGGGCTTATTGGTATTAAAAACATTGATCTGCCAATGAAAGTTCGGCT
CAACACCAGGAAGAAGCAGTCCCGCCAAAATAAAAGGGTCCTGGGCCGGAGTA
TTGAAGAACGCCCGGCCGAAGTCAACGACCGACAGGAATTCGGCCACTGGGAG
ATCGATACCGTCATTGGCAAAAAGTCGAACGACCAGGCCTTGTTGACCATTACTG
AACGCAAGACACGCAAGGAGATCATCATAAGGGTCGATGCAAAGGATGCGCCG
TCCATTAGTGATGCCATTTCGTGCTTGAAGAAGAACTATGGTGAGCAGTTCCCGA
AAGTGTTCAAAACCATCACTGCTGACAACGGGTCAGAATTTGCCGACCAAACCA
TTGCCTATGAGGAAAGCTGGTGCAACCGCCTTCCAAGGAGGATCCTGGGATACA

GGACACCGGATGAATGTTTCAAGGAGGAGCTGGCCCTGTTGGCCTCTTAAACTC
ATCAGCGCGGGTCAGTTTGTTGAGTGTTTAGCAACTTGACATGGAGCCTCCGCAG
GCATGATTCTCCTGCCGAGAAGCCAGATGTCCCAACATCGGCATCGCCTAACGA
GGCTATCATGCCTGCCACTCCATATAAAGTTGCGCGGAGTGAGTCCTATATCAGT
TAATTTTCATGTTCGACAACGGTTTACAGGGTTATCAGCATGTGTTGCATTTAAT
ATTGCAATTTAGGATAGTTTAATAGTTTAAAAGGTTAGAGGAAAAAAAGTGAGT
TTGTGTAATTGAAAATTTGGCTTTTTAATTGAGAATGAAAGGAGGGAGAAATAT
AAGACTTTGTGCAAAGTAGATACTAGTGGTATGTTAACACTAAACTAAACAATTT
ATTGGGGTTTTGTTAAAACAATACTGTTACTTCCGCCAGTATCAACAATTTCTGT
ATCATTTATTGTTCCTAATACAAATCGATCTGATAAATCTAGAGTACCATTGGTA
CCATCACACAATAACCACTCAAAAGGAATATTATTAATTGACTCATTCCACATAA
CAATTACTCCTCTAGGGAGTATTCCTTCGGGACCCCTTGGTCTTCGTCGTCCAAGT
CGGTTATTAACCATATAAATCACCTCATAAGATTTTAAAAAATTACCTTATAAAA
TAATATGTTAGTGATTTAAAAAAACATTGGACATTTGTTAACTTTCCATTATTACC
CTAACATACATTATTTGCTAATTCGTTTTGGTATCTTATCGAATCAGAATTTGTCA
AAAGGCGACTTAAATCGCTTTAGCGTATAATATCCATGGAATAATTTATGTATAT
CCTATACAAATTCTAGTTTAATAAAGCATAAAAAAGAAAGGTTTTATGCAATCCC
TTCAACTAACGTACCCGTTCGCATGTTTGTTATAGAGAAGTCTGTTAAAGATTAG
GAAATTTATAATCAGGATATTGAAGATCAAATTGCCCACACTGAAAAGTCCCAG
TGGTTTCAGGAGCGTTCAAATCGAGGTATTGCAAATCATTTGCCCCAGGGTGAAT
GATCTGGAATCATGAGTAAACCAGACGGCTCTTTAATCGGGTCGTTTTTATTTTT
CAACATAATGTCAACTATGTTGCGTTTTATCGGTAGAAATGGTATTTTTTAGACA
ATAAATTCAACATAGTTTATCAAAAGAGTACAAAAAATAAATTCAACATAATTT
ATAAAGGAGAAGAAATAATGATTGATGAATTTAAAAAGTATTTACTGGAAGAAG
GGAAAAGTGAACATACAATTAAAGGGTATATCCAGTCTGTCAGCGGTTTTTTAA
AGTGGTTCAGTGAGTCAAAAGACGTAGAATTTACAAAGTTGCATCGTGAAAACG
TAAAAGAATACATTTCTTTTCTTCGAACCGTAAAAAAAGCAAAACCAAAAACAA
TTAACACTAAACTGAATGCGTTAGTTAAATTCAATAAATTTTTAGTAGAAACGAA
CATTCAACGGGATATGATCTTAACAAAAAAGGATTACATGAAGGCGCAGCAACA
ATATGCTTCGTTAGCCAAGGTGAAACTAAAGGACGTCGAGAAATTTCGGCAGTT
GGTGTTAGATAGTGGAAACAAACGAAATTATGCATTGGTTACTCTTTTAGCTTAT
GCTGGGTTGCGAATTTCTGAGGCATTAAATCTAAAAATGCAGGATTTCAACCTAG
TTTCTAGAGAAATCACTGTTCGAAGTGGTAAGGGAGATAAGTTTAGGGAAGTAT
TTATGAATGACAAAGTAAAAGCTGCTCTCCAGTCATGGTTAAAGGAACGAAAAG
AAAAAAAAATCGAGCATGATTTTCTTTTTATTAGTAATCGTAATCAACCTTTAAA
CCGCACAACGATCAATAAATTGTTTAAAGAGTATTCGGAACCTATTGGTAAAGA
AATTACCCCGCACGATTTACGGCATTTTTTTTGTTCTCATGCGATTAGTAGAGGAT
TGAGCGTTCATGAAGTAGCGAATCAGGCGGGACATTCCAATATTCACACTACGC
TCTTATATACAAATCCGACAAAAGATGAACTTATTCAAAAAATGAATCAACTAT
AAAAGAGAGAATAATGCTTTCTTTTTTCCTGTTGATTTTAGATAGAATTGTTTATA
AAGGATTTGATATGATTGATATCGAATACGTATCATATGGAGGGGAGGAAAAAT
GAAGGAATACTCTACAAAAGATATATCCAATATTGTGGGCATAGCGACTCCAAC
TGTACGAAAGTATGCCCAAGCTCTCGAAAAGGAAGGGTACATATTTATTAGAAA

TGAAAATGGCTTTCGGATATTCGTTGATGAGGATATCGAGATACTCCAACAAATG
AAGCAAATGTCAAACGAATCGGGCATGAATATTGTTCATATTGCATCTATACTTA
TAAATCAACGAAAACAAAATCTTGTCGATACGATACAAAATGAATCGGAAGTCA
CTACATCTGTAAAAAGTGAAGTAAAAGAAGCGGAATCTCTTAATATAGATCGTA
TCGATAAGAAATATGATATGTTATTGAAGGAAGTACAGGAACTAAAGCAACTCG
TGAGGGAACAACAAGCATATATTGACGAGCGTTTAGAGAAAAGAGATCAAATG
ATTACGCAATCCATTCGGTCGCTTCAAGAACAACGGCAAGCTTTGCTTGAAGCAA
GTACGTCCAAAGAGAAAAAAGGATTTTTTGCGAGATTGTTTTCAAAAAGACTAT
AAAGATGTAACCGCCAAATAGAAATGAACACTTTCTTACCTAAAAATGCTGAAT
AAGATTGAACACTTTTTTCCAGTAGAGAAGTGCACTTTATTTGGAGAGGCGGGCA
TGATAGTCTTTGTCGGCTGAGCCAGATTCATTGATATATCTGGCTTCCTGGCTAAT
AATCTTGCCCGCAGAGGCTCCAAGTAAAGTGTTCACTACTAATTAGCAAAAAGT
GTTCAAAGTTATCAGGCGGTTACATAAAGAATATAGTAAAACTTTTTGTTACTCA
TCCGTCCTAGTAATAGTAGTGTATTTACTAATACTATATAACAAAAAAGTGCGTA
CTTACAAATAAAAGCTATAGATATTATACTTGTATCTATATAGAGGATTATCACC
ATAAGAAATAAAGGTGCTTGCAACGGACTTTTTACATGGTTGGAAGAGATCCTG
AAGATTGGAGCATGATATAGCCACTTTTTCGAGATACAGCTGTAGAAAAAGGAT
ATTTATATGCAGAAAAATTCATAAAAGAAAGCCGAGGTATTAAGTTATGATCAA
GATTGCACCTTCCATTCTTTCAGCTAATTTTGCGCGACTTGAAGAAGAAATAAAA
GATGTAGAACGGGGCGGAGCCGATTACATCCATGTTGATGTCATGGATGGTCAT
TTTGTGCCAAATATAACAATTGGTCCATTAATTGTCGAGGCAATTAGACCTGTCA
CAAATTTACCTTTAGATGTGCATTTAATGATAGAAAATCCAGATCAATACATTGG
GACGTTCGCCAAAGCAGGAGCCGATATATTATCTGTTCATGTTGAAACTTGTCTT
CATTTACACAGAACCATTCAATATATTAAATCTGAAGGTATAAAAGCTGGAGTG
GTATTAAATCCCCATACTCCCATCTCAATGATCGAGCACGTGATAGAAGATGTTG
ATCTTGTATTGCTTATGACGGTTAATCCAGGCTTTGGGGGACAATCTTTTATTTAT
TCTGTCCTACCCAAAATAAAACAAATTGCTACTATTGTAAATGAAAGAAAGTTGC
AGGTTGAAATTGAAGTAGATGGTGGAGTAAATCCTGAAACGGCTAAACTTTGCA
TAGAAGCAGGAGCCAATGTCCTTGTTGCAGGTTCAGCCATATATAATCAAAAGG
ATAGAAGTCAAGCCATTGCAAGAATTAGAAATTGAAAATAAAGACGTGGTATCA
AAAAAAATTGGAAACTATACCTTTCAATAGATTCTTTTTTGTCTTTTATAGTGATG
ATGAAAATTTATTCGGTGCGTTGATGACCGAAGTCTTCTTTAGACAATTTTAACA
ACATCTGAAATTAGGAAAAAAGGGCAACTTTTTGTATTTTTTGTTGCCCTTTTTCT
AGTGTCATTTTATTTAGATCAGTCCTATTCCTACTATAGAGAAAGGCTTCTTGCTA
AATAATAAATGGACAAATCAATCGTATGTTTATCTTTCGCTGCTTCCTCAAAAGG
TGTATGAATCAGTTTACCATTTCTCAATCCTACCATCAGACCTTTTTTGCCTTCAA
CCAGTAAATCCACGGCTTTTGCACCCATCTGGCTGCTCATCATTCGGTTATAGGC
GCTAGGAGATCCCCCACGTTGAAGATGTCCAAGTATGGTCACCTTTGTGTATCAA
ATCCTGTTTTTTCTTTAATCGCTCTGCCGATTTCTACAGCTCCTACTCCATTAAAT
GCTCCTTCAGCAACCACAATAATGCTGTGCGTCTTTCCACGCTGATATCCTTGTTT
AATACGGTCAATTACATCTTCCACATCATGGTCTGCTTCTGGGATAATAATCGAT
TCTGCTCCCGCACACATTCCTGCCCATAGAGCAATGTCTCCTGCATCGCGACCCA
TTACTTCAACGACGTAGATACGCTCATGAGAGGCTGCCGTATCACGAATTTTATC

AATTGCTTCCACAGCTGTGTTCACAGCAGTATCAAATCCAATCGTATATTCCGTT
CCTGCAATGTCATTGTCAATGGTTGCCGGGATACCAATAGTTGGAAATCCTTGGG
CAGTTAATTTTCTTGCACCTTCAAAAGTGCCATCTCCACCAATCACGATTAAGCC
ATCAATGCCCTCTTTTTTTAGTTGTGCTAAAGCCTGTTGTTGCCCTTCAACTGTCT
TAAACTCTTTACATCGTGTAGTTTGTAAGATAGTCCCTCCTCGGTGAATAATATC
ACCAACACTTCCAAGGCTCATTGAAATAAAATTACCATCCATTAAACCTTTGTAA
CCATTTTTTATGCCATAAACATCAAGTCCCTTAAAAATCCCTCTTCGAACGACCG
CACGAATAACAGCGTTCATTCCTGGTGCATCCCCGCCGCTAGTTAATATTGCAAT
CTTATTCATTCTTGTTTTCCTCCATATTTTCTACTTCTTATATACAATCCACGATGT
TTATTTGCAGTGCCCAGTTAAAACAACACTTTTTTTGGCATTGATGAAACCAGTT
TATCTAAGAGGAAAGGAAAGATATCGATTTGGTCATGTATAATTAGACTGTGTC
GTAGTAATCGGAGTTTATATGCATACATAAAAACAGAAAAAAAAGAGGACGTTA
TTGAGATATTGAAACAATTATCCCGTCTTGAACGCCTCATCGGAATTGTGATTGA
TATGTGGAAACCCTAAAAAACAAACTGTTTGGCTTCGTTTTTTTACATCTAAACT
CCAAAGAAATAAATGAATTTATGTTTCTAGTGAATATTTTTGCGAAAGAATATTC
TAATCTTTCTAACATTTAATGAAATACTAGTTTCTTACCGATTAAACAAGACAAG
AATTTTAGAGAAGCTTTTTAAAATGAGAAACGACATTTTCGACCGTAAAACCGTA
TTCTTCAATTATTTTATCTCCTGGTGCTGAGGCTCCAAATTGATCAATGGCGAGG
ACGTTACCGTTATCACCAACATATTTACTCCAACCTAAGGAAGCCCCCATCTCAA
TACCAAGACGTGCTTTTACGTTTTTTGGAAGTATACTTTCTTTGTATGCATCAGAT
TGTTTTTCAAAACGATCCCAGCTAGGCATACTAACAACTGAAACATAAATACCTT
CCTTTTCTAACTCTGCTTGTGCTTCAACAGCTAAAGCAACTTCAGAACCAGATGC
TAATAATAAACCAGCAACTTCACCTTTTGCTTCAGAGATTACATATGCACCTTTTT
TAACACCCTCATACGCCTTTTCAGAATCAACAAGTGTTGGTAAATCTTGACGTGA
GAGGATAAGAGTTGTTGGTTCGTCTTTACTTTCTAACGCCAACTTCCAAGCTGCA
GCTGTCTCATTGCCATCAGCCGGGCGAATTGTAGAGATACCAGGCATTGCACGTA
AGGATGCCAATTGTTCAATTGGTTCATGTGTTGGTCCATCTTCACCTACAGCAAC
GCTATCATGTGTAAAGACATAAATAACAGGTAGTTTCATTAATGCTGAGAGACG
AATGGCCGGACGTAAATAATCAGAGAATACAAAGAAAGTTGCTCCAAATACTTT
TACACCACCGTGTAGAGCCATTCCGTTAACAGCAGCACCCATTCCAAATTCACGA
ACACCAAACCAAATATTACGTCCGCTGTAATCTGCAACGCTGAAATTTGCTTCCC
CTTTTAGCAATGTTTTATTAGATGAAGCTAAATCCGCGGAACCGCCAAGTAGTTG
CGGAACATTTTTCGCTAGGGCATTTAACACTTCACCACTGGAAGAACGAGTAGC
AAGTTTATGTTCTCCGACACGGTACACAGGAACATCAGCATCCCAGCCTTCTGGA
AGCTCACCATTAATCGCTTGTTCTAATTCCTTTGCAAGTGCTGGATATGCTTCTTT
ATATTGTGCGAACAACTCATTCCATGCTTGTTCCTTCTTCTTTGCCGCTTCTTTTA
ATTCTTTAAAATGTTCTTTTACTTCTTCAGGGATATAAAAATCTTCTTCATACTTC
CAGTTATAATGTTCTTTTACAAGCTTTATTTCGTCTTTTCCAAGTGGTGAGCCGTG
CGAATCAGACTTGCCACCTTTATTTGGTGAACCATATCCAATTATTGTTTTGACTT
CAATTAATGTTGGTTGGTCTTCGTTGTTTTTCGCTTCAGCTATCGCTTTTGCGATT
GAATCGATGTCATTGCCGTCCTCAACACGAAGTACTTGCCAGCCGTATGCTTTAA
AACGATCTTGAACACTCTCACTAAATGACATATGAAGATCGCCATCAAGAGAAA
TATCATTTGAATCGTATAATACAATTAAACGACCAAGTTTAAGGTGTCCTGCAAG

TGAAGCTGCTTCTGCAGATACACCTTCCATCAAGTCACCATCGCCACAAATGCTG
TATGTGTAGTGATCAATGATATTAAATTTCTCACGATTGTATTTAGACGCTAAGT
GTCTTTCAGCCATCGCCATCCCAACTGCCATGGCAATACCTTGCCCTAACGGACC
TGTTGTGGCATCAACCCCAGGTGTATGGCCAAACTCAGGGTGACCAGGTGTTTTG
CTTCCCCATTGGCGGAAGTTTTTCAAATCTTCTAATGATAGATCATAACCAGTTA
AATGTAATAGGCTGTATAATAACATGGATCCGTGGCCCGCTGACAATACAAAAC
GGTCACGATTAAACCAATTCGGGTTGTTTGGATTGTAATTCATAAATTTTGTCCA
AAGTGTATAAGCCATTGGGGCAGCCCCCATTGGCATCCCCGGATGACCTGATCC
AACCTTTTCAATTGCATCGATTGATAGAGTTCTAATAGTTTGAATGGATAACTGA
TCAATATTTATTTTTTGTTGGAGCACGATTTCCCCTCCCCTAATGTCTGTTTTTATT
AAGCTTTACCTGAAGACCCAAATTCACGCATTTTACCTTTAATAGTTTCTTTAATC
GCTTCCCGTGCCGGTGTTAAAAATTTGCGAGGATCAAATAGCTTAGCATCGTTGT
TTAAAACTTCTCGAACGGCTTTTGTAGCAGCAATTTGGCTTTCTGTATTTACGTTA
ATTTTTGCTGTACCAAGCGAAATAGCTTTTTGAATATCTTTTGTTGGAATACCTGT
ACCACCGTGGAGAACAAGAGGAACACCTGTTAATTCCATAATTTCTTTCATCCGA
TCAAAACCAAGGTTTGGTTCACCTTTATACGGACCATGGACAGAACCTAACGCA
GGTGCAAAGCAGTCTACTCCTGTTTCACGAACTAATTGCTCACATTCTGAAGGGA
CAGCATAAAATGATTCAGCTACTACATCATCTTCTTGTCCACCGATACGACCTAG
CTCTGCTTCAACAGATACACCAACAGCATGCGCTATTTCAACCACACGTTTTGTT
AATTCAATATTTTCTTCAAGTGGAAGATGGGAACCATCGATCATAACAGATGTAA
ATCCAGCATGGATGGCTTGTACACATTTCTCAAGGCTTGGACCATGGTCAAGATG
AATAGCAACTGGTACCGTTACGTTATATGCATCCATTAATGATTTGACCATATCG
ACAATTAACTTAAATCCACCCATGTAATTAGCAGCACCTACAGATACCCCAATA
ATAACAGGAGACTTCTCTTCTTCTGCAGCTTGTAAAATCGCTTGACCAAACTCAA
GATTATTGATGTTAAACTGACCGACAGCATATCCATTTTCTTTTCCACGATTTAAC
ATATCCTTCATTGAAACTAATGGCATAATTTCCACTCCTTATATAGTATTTTTAGA
ATCCTCTTTCGAGGTAATTTTCAGATCATTTATTGCTCTTTGATAACTTTATAAAG
GATCGTTCAGATAGTGGAGAGCTAAGCTGTCTAATATCTGAACCATCCTATCAAT
AGGGGTTGACTTTTGATAATATCCCTAGATTATGGTCTTAGGATCGTTTTGGGGA
TTTCTTACCACCTATGTTTTTTCATGAATTTCACGAAACGGACCATGTTAATGCTT
AATTATTGCAGATCAACTACGTTTTGACCTTCTTCAATTTTCTGTTGTCACGCAAG
AACAATTTATTACCATTCAGAGATTAACAAGGATAAGATACCATCTTGTTCACAA
CAAAAGAAAATATTCCTTTAAATATCTGTTTTACAAGTGTATTTCCTTTACGGAT
AATCAGGAAAGTTTCAAGCTGAGGAAACTAACATGATAAGAGCAGGAAATCACT
ATCTTCGTTATTACCTAATTAAAGTTGCTAACTCGTTATAACGACATGAACCGAA
TTTCGAGAAATTACATAAAAAAGTATCGTGAGGTTACGAAACATCAATATAAAA
GAGCACTCGATTAGATACATTAAACGTCATCTATTACGTTTATTATCAATGATAA
GCTTCAATAAATTGGTATTTTACGCTACTTCCTCCAATTGTAAGAAAAGAGTGTG
TTTGCATTAAACCATCTTCTTTTTTTCGAATACCCTTCATAAGCAGTCCGTCAGTA
ATTCCTGTTGCTACAAAAAAGCAATCATCTGATTTTACAATCTCATCTATTGTAA
AAATTTTTCTTGGATCTTGTATTCCCATCGTAATACACCGATCAAATTCTTCTTCG
TTTTGAGGAGCAAGTCTTCCTTGGAAATCTCCTCCTAAACACTTCAGTGCGGTTG
CAGCAATTACTCCCTCTGGAGCACCGCCAGTCCCTACTAATATATCAACGTCTAC

TTCTTCGAGTGCAGTACCTATTGCCCTTGTAACGTCACCATCAGAAAATAATTTA
AGTTTAGCCCCTACATCTCTTACTTGTTTGATCAAATGATCATGACGTGGTCTATC
CTGTATCATTACTGTTAATTCTTTCAAATCTTTCCCTAAAGCATTAGCAACTGATT
TCATATTTTCTGTTAAAGATGCGTCTAGATTTACACAGCCTTTTGCTTTTGGTCCT
ACAGCTATTTTTTCCATGTACATATCTGGGGCATGCAGTAAACTTCCTCTAACGG
ATGCAGCAATTACTGCTATTGAATTATCCATTCCTTTTGCCATTAAACCCGTTCCA
TCAACAGGATCTACTGCAATATCGAGTTGGGGTCCTTTTCCTGTTCCGAGTTCCTC
TCCAATATAAAGCATAGGAGCTTCGTCCATTTCCCCTTCACCAATAACAATTAAA
CCGCTCATATCAATGAGGTTCAGTCGATTGCGCATAGCTTCCGTACCAGCCCTAT
CAACTTCGTTTTTATTCCCTTTTCCTATCCAAGGATAAGAAGCGAGAGCTGCTTG
CTGTGCTACAGAGAGAAAATCCATAGCTAACGCCTGAACACGTTTTTCGCTTTTC
AATTCACTCATTTGATCACCTCCTGTGTTATTTATACTTTCTATATTCGATATCAC
CTATACAAAATCTAGTATAACGTGTATAATATCGATTCGTAAGTATGCACTTTAT
TAACATATAGTATCGAATAGTATACTGTATTAAACGAGAGAGTTTTCTAATCGAT
GGACTTAGCTCTTTTTATTTCGTTTTCAAGACCGTTTTCCATGTAGTTTTCCCCATT
CACACACTCTTTATTTGATTTCGTAACGGGAAAATTTGGTTTAAGACATTTTCATG
GACCGCAATGTTTTTTACGATACGAGACTTCATTCCTTAATCTTTCATTATGCGGG
AGACCGTCTTCTGAAAAACCTCAATTCCTTCATGATTGAGCTTTTTCATAACTTTT
GAACTCCTATACAGTTGGCGGGATTCGATATATTTTTAAAAATCTTCCTCGTACC
GCCTTTGTTTTTCGACATGTAGTCTCCTCATTCTTTAGTCTATTACTCTATTAATCT
ATTAATCTGTGTCTACTATATTAATTTAGCATCACGAAAAAGCGTCTCTTGAAAT
TCAAGAGACGCTTTTCCAAGATTATTATTACAGAGCGTTTTTGATGATTTGTGCA
ATGTCTTGAACAGTAGGAACACGTGGGTTATCTAGAGTACATACGTCTTGGTACG
CGTTGTTCGCTAATAATTCGATATCCTCTTCTTTTACGCCCATTTCTGCATAGCCA
GATGGGATACCGAAGTTTTTGTTATAGCGTTGAAGCGCTACAATTGCTCTCTCAG
CAGCAGATGCAGTGCTTAAGCCAGAAACATTCTCGCCTAAAAGCTCAGCAATGT
GTGCGAAGCGTTCAGTACGAGCAATTAAGTTGAATTGGCAAACGTGTGGCATAT
TAACTGAGTTACAGATTCCGTGTTGTAACTTGTAAACTCCACCTACTTGGTGAGA
AATAGAGTGTACTAATCCTAAACCTCCGTTGTTAAATGCCACTCCTGCCATGTAT
TGTGCATAAGCCATTGCTTCACGTGCTTCGATGTCTTCTCCATTTGCAACCGCACG
TGGTAAGTATTCATTAATGAGTTTCATTGCTTGAATTGCAAACGCATCAGTAACT
GGTGTAGCACGTTTTGCAACATATGCTTCAATTGCATGGGATAATGCATCCATAC
CAGTTGCAATTGTTAATCCAGCTGGTTTTTTCACCATTAATTCTGGGTCAACAATT
GCTACAGTTGGTGTAATTTTCTCATCGATAACTGGCATTTTTACTTTACGTGCAGA
ATCTGTAATAACCGCAAGAGATGTTGTTTCACTACCAGTACCAGCTGTTGTAGTG
ATTGCAACTACTGGAACAACCGGTTTTTCTACACTGTTTACACCTTGATAGTCGTT
GATTCTTCCGCCGTTTGCTGCAACTAAACCGATTGCTTTTGCTGTATCGTGAGAG
CTACCTCCACCGATAGAAACAAGTGCATCACATTTTTCTTGTTTGAATATATCTA
CGCCTTCATGAACTTGTGTATCTGCTGGATCTGGTTGAGCTTTTGGGAAAATTAC
AGCATCAAGGCCAGCTTCACGAATGTTTTTAGCAACTTCTTCTGACAAACCTGTG
CCATGAAGAAATGCATCTGTAACGATAAGTGCTTTTGTAGCTCCAATTTGCTTAA
GTCTTGTTCCTACTTCTTTTACAGCGCCGCGTCCAATTACGCTAGCTGGTGGAATG
AAAAAGTTTCTTTGCGTCATTTTGACTACCTCCTAATGTATGTAATTGTTTACAAA

ACTAGTTTAGATTCTAATCTAGGTGTTTGTCAAGATAGTTGTCGTAATGTTCGTAT
TTTTCGAAATAGTATTTTTTACAAGGAGCTACTAGATTACTGTTGGTATGAGGCT
TTCCCCTTAGGGAGAGTGTCAATTTTAATAAAAAGCTCAAGGGAAAAACCAACC
TAATAAAAGGTTTCAAAAAGGAAAAGTGATTTTGACGCTTATATGGTGTGGAGG
GTTAAAATGGAAGGGCAGTATTCAATTTGACTTTAAATAGGATTAAAGTCATATC
ATAAAACGACTGATCGACTAGATCCCAAGGCATCAAATTGTCTCAAAAAAATGA
ATGTTTTTTGCGGTCAACTTGAAGCTTTTATTTCATATCCGAATGATCGTTCTATC
ATGATAATAGAATAGGTAATAGATATAGATAAACTCAAGGGTAGTCAGTATTCC
CGAAAGAGGGAGCTACTTATACCGTTTTTCAAGCATTGATGTTCACTATTTCGTT
TTCAAATCAAATCGTATAATATATTGCTTTAACACAAAAAAAGTAAATCCATCCT
TGAGCCGGCAAGCGTTAAGGGTCTGGATTTACTTTTCTTTGGTTAGATGATCTCC
CTTTGAGGGCTTTAACTCCGTAGACATTCAACAACCTCCTTCAATTTATTTCTTTT
TTTACTTTACTTTTAAGTTAACTAATCTTTTGAAAATATGCGAGAGACTTTCTTTA
CAAGTTGCAGAAGCTGCTTTAAGCCAGGATATTCTTGTTATTAGCTAATTTTTAC
CATTTTTTCCATCTTTTACTAAATTCTTTTTTACGTGTTATAAAACTCTTTTTAAGA
TATGTAAAATTGGTAAGTAATCTACAATAAACCTTTATCAGTAAATATGGTCCTT
AAAGCTTCGTTTACAATTTTAGACTTTGCTCCTCTTCCACCCTTTTTGGCTAAACG
GTCCAAAACTTGAGATAAATCTTTTTGAAGATAAATTCCAGTTAACACTGTCTCG
TTCTCTTTCTTTTTCATGTTACCCTCAATCAACGTATCCAAGTAATCTTCTTGCTTT
TTAGTACCTGAAGAAACATTTACGTCATTATTAATATTAATGTGATTATTATCTTC
GATATTAATATATTTATTAATATCATTATTAACATTTTTGTTTATGCTTTTAATATT
GTTACCATTAACTTTATTATCATTATTTTTATTAATATTATTGTTAGCAAGATCAT
CAAAACTTGCTAGGTTTCTATTTTTTCTTGCCATTATTCTTGTACCTCCTGTAATA
AATTAAAATAAGCTTGTACCAATGGGTTTTTAAAATCAGTCAGTGTTGCCGGTTT
ACGTTCGTAAGCAACGCTTGCAGCAAATCTTACAGAACGAGGAATAACTGTTTC
AAACATTCTGATACCATTTTCCGCACAAAATCTTCTGCACTTTTGAAGGACCTCT
GAGTGTAGGGTAGTTCTTTGATCTACAAGAGTAGCGATTACTCCCTTAATTTTTA
ATTCCGGATTATGCTGTTCTTTAAAATTATGTATGGCACTTAAGATTTTGATTAAA
GACCTCATGCTATATCCTTCAGGTTGGAACGGGATTAAAACACTTTCCGCATACG
AGAGGATGTTTCCTTGAACTAACCCTAAGTTTGGTGGACTATCGATTAAAATATA
GTCGTATTCTTTTTTAATTGCCTTCATAGCATTTTTAAGCATTTTAAATGGATTCG
GGTATTTTTCTCTATTAGAAAGAACATCAAACTCCAAAAAGCTCATATCATCATT
TGCCGGCAAAACATCAATATTGGGATGAACATTAATAATTGCCTTTTTTGGATTT
AATCCTTCAACTAAGACATCATATAAAGTTTGTTCTACTGAATCAGGATTAATCC
CAAACGAAACTAATACGTTTCCTTGGTTATCCGTATCGATAATTAAAACTTTATT
GTTATTGCATAAAGCTCCGGCTAGATTCGTAGTTATGCTGGTCTTTAAAACCCCA
CCTTTGTTAGTAGATACTGCAATGATTGGCATGTTTTTATCCTCCATTAACATTAA
TATTAATGTTAATTTTAACACGTGTAATAGAAAAAAAGAAACTCTTTGTAGACTA
CATGGCTTGTTATTAAATCTCATCTTAAATATAAACTATTGAAAGTCAATATGTC
TGCAGATAAGATTTCATAAACTTTAATTTATCTATAAAAATTTAGTCCACTCCATT
TTTAATTACTTTATAGGCTTGGTCGAAATGAATAGTTTTTCGGAAGTCTAAACCA
AAACGATCTGTTTTTCCTATCTGGTATATTGAATTTTCGCCTTTAATATGTTTTTCT
TTCTCAAGAAAAGAAAATAATGACATTAACTAAATTCATAATTGCTTGTTTAATA

TCCTTAACTTTAGATATTATTCTACTTCCTTTAGCTCAACATAAAAAAGAGCCAA
TCTTAGGCTCTTTAAATTAACCAAAAATGTTATAAATTTAAGCATCAATTGTTAA
TTTCTCAATCCTTTTATACCTTCTTCTAATTCCTTCTTTTTCTTTAGAAATTCTTCA
TCAGATAGTTCATCAACTTCCGATTCCTTTATTTCTAATACATTAAACAATTCAGT
TTTAAACTCATTGAATTTCGTCAAAGCCTCATTTTTTTCCATATCGAATTGATTCT
GTATTTCTTCTATAGCTTTTTGTTTGACAAACCAATCAGGCTTAGGCTCTTTCCAT
TTCTGAAAATAGGAGATTAAATGAGTAAGAATCGTTTGATCTTCTGAAGTTGTAG
CCACCTTATTAGAATTATTATTTGCAGAAGTATTTAATACAGTTGTTATGTATCCG
ATCGGGTTTCCTATGTCTGTTCTTCCTTGGATCTTCTCCATTAAAAGAAGAACGTT
TTCTTGTCCGTATTTTTCCCACGATTTAACAAGTTCATCAGACACTTGAACACCCA
TTTTTAAGCCAAATTTTTTGACCTGTTCAGCAAAGGTATTTGGTAGTTGAAAATCT
TCTAAATTCTCTTCAAAAAGCGATAATTGCGGGTTGTTCTTTTTCTTAGTCGGAAA
TATCAAAAATTTAATTTTATTAACTCGGCGACCCACTTTAATTTCTTCAATTTCAA
AGGAGATATCTGTCTTTTTCTTTAATTCTTTTTGGGCCGGCACCAGAACGCGCTGT
TTAAAATTTCCATACGCCGGATAGATATCTTCGGCGCCAAGCATTTTGCGTAAAT
CATCAAGAAAAAAGGTTCGTTTTTGAAGCTTTTCATACTGTTTTAAAAGTTCATA
TATACGTATGGCATAAGAACTTTTCAGTTTAACAACGTTTTCTAACTTATAGCTC
GTAAACTCTTTTTTTAATTCCAAAAGATATGGTCGTAAAAAAGGATCGAACCTAA
TATCAATGGTTCCTTCAGATTCGTTATAAGCAACATAACTCAGCCAAGCAACCTG
TATTACCTTTTTGTTAATACGAACTTCAAATACTTTCTGCATCAGTTCTTTTGTAA
TTTTCCTTAATTCTGTATACTTAGGAGAACCTTTGAGGCCTAGCAATTTATTAAAC
TCTTTAACGGGTAGCGTATATGTTTTGAAGTCAGAATCAGTCGGCCGAATGTTAC
TGGCCAGACAAAGGATAATCTTCTGTTCCACTACCCCCAATTTATAGTTTGCTTC
AATAAGCATATTTGACTTCGTTACAAAATTACTGGTCCTTGTTTCAATTAATTCTG
TTGCCAACACAAAGATTTCCCCCCTTCCCATAGACTAATTATGACAAATTCTGAA
CTATAACGGAAGTTTTTTCTTGTTATAGATGACTTTAAAATAAACAAAGCACCAT
TTTTGTTATACTTCAAACCAAAAATGTTATATATGAACCATTTTTGTTATAGTTAT
CACCATTTTTGTTATAGTTAAGGAACCAAAAAAGTTATAGTTAACACCATTTTTG
TTATATTTCAAACCAATTAAGTTATAGTTAACACCAAAAATGTTATATTTA

APPENDIX B
Annotation of *B. methanoliucus* MGA3 genome

```
     CDS             complement(1..772)
                     /note="site-specific recombinase/integrase"
                     /gene="BMETHMGA300001c"
     fasta_record    1..19105
                     /note="ConsensusfromContig1 Average coverage:
            2,780.81"
                     /label=ConsensusfromContig1 Average coverage:
            2,780.81
                     /colour=10
     CDS             complement(1533..2333)
                     /note="hypothetical protein"
                     /gene="BMETHMGA300002c"
     CDS             3645..4637
                     /note="replication initiator protein, RepB"
                     /gene="BMETHMGA300003"
     CDS             4910..5749
                     /note="partition protein ATPase"
                     /gene="BMETHMGA300004"
     CDS             5749..6165
                     /note="hypothetical protein"
                     /gene="BMETHMGA300005"
     CDS             7252..8400
                     /note="NAD-dependent methanol dehydrogenase"
                     /gene="BMETHMGA300006"
     CDS             8949..9911
                     /note="fructose 1,6-biphosphatase II"
                     /gene="BMETHMGA300007"
     CDS             10505..11359
                     /note="fructose-1,6-biphosphate aldolase"
/gene="BMETHMGA300008"
     CDS             11375..13393
                     /note="transketolase"
/gene="BMETHMGA300009"
     CDS             14047..15015
                     /note="6-phosphofructokinase"
                     /gene="BMETHMGA300010"
     CDS             complement(15296..15892)
                     /note="hypothetical protein"
                     /gene="BMETHMGA300011c"
     CDS             complement(16147..16521)
                     /note="transposase"
                     /gene="BMETHMGA300012c"
     CDS             complement(16528..16845)
                     /note="transposase"
                     /gene="BMETHMGA300013c"
     CDS             complement(17505..18149)
                     /note="ribulose-phosphate 3-epimerase"
                     /gene="BMETHMGA300014c"
     CDS             complement(18432..18977)
```

```
                /note="hypothetical protein"
                /gene="BMETHMGA300015c"
fasta_record    19106..24199
                /note="ConsensusfromContig2 Average coverage: 309.89"
                /label=ConsensusfromContig2 Average coverage: 309.89
                /colour=11
CDS             complement(19250..19978)
                /note="ABC-2 type transporter"
                /gene="BMETHMGA300016c"
CDS             complement(20031..20834)
                /note="sulfate-transporting ATPase"
                /gene="BMETHMGA300017c"
CDS             complement(20836..21273)
                /note="MarR family transcriptional regulator"
                /gene="BMETHMGA300018c"
CDS             complement(21926..22654)
                /note="ABC-2 type transporter"
                /gene="BMETHMGA300019c"
CDS             complement(22707..23510)
                /note="sulfate-transporting ATPase"
                /gene="BMETHMGA300020c"
CDS             complement(23512..23949)
                /note="MarR family transcriptional regulator"
                /gene="BMETHMGA300021c"
fasta_record    24200..26742
                /note="ConsensusfromContig3 Average coverage: 477.50"
                /label=ConsensusfromContig3 Average coverage: 477.50
                /colour=10
fasta_record    26743..64147
                /note="ConsensusfromContig4 Average coverage: 644.62"
                /label=ConsensusfromContig4 Average coverage: 644.62
                /colour=11
CDS             complement(26828..27919)
                /note="metal dependent phosphohydrolase"
                /gene="BMETHMGA300022c"
CDS             complement(28026..30554)
                /note="DNA gyrase Subunit A"
                /gene="BMETHMGA300023c"
CDS             complement(30754..32676)
                /note="DNA gyrase subunit B"
                /gene="BMETHMGA300024c"
CDS             complement(33028..34146)
                /note="DNA replication and repair protein RecF"
                /gene="BMETHMGA300025c"
CDS             complement(34572..35708)
                /note="DNA polymerase III subunit beta"
                /gene="BMETHMGA300026c"
CDS             complement(35886..36902)
                /note="Cromosomal replication initiation protein
       DnaA"
                /gene="BMETHMGA300027c"
CDS             38057..38425
                /note="ribonuclease P"
```

```
                     /gene="BMETHMGA300028"
     CDS             38535..39347
                     /note="OxaA-like protein precursor"
                     /gene="BMETHMGA300029"
     CDS             39344..39964
                     /note="SpoIIIJ-associated protein"
                     /gene="BMETHMGA300030"
     CDS             40390..41775
                     /note="tRNA modification GTPase"
                     /gene="BMETHMGA300031"
     CDS             41979..43868
                     /note="Glucose-inhibited division protein A"
                     /gene="BMETHMGA300032"
     CDS             43896..44612
                     /note="Glucose inhibited division protein B"
                     /gene="BMETHMGA300033"
     CDS             44729..45598
                     /note="ParB-like partition protein"
                     /gene="BMETHMGA300034"
     CDS             45788..46567
                     /note="chromosome partitioning protein parA"
                     /gene="BMETHMGA300035"
     CDS             46560..47417
                     /note="parB-like partition protein"
                     /gene="BMETHMGA300036"
     CDS             47395..48168
                     /note="hypothetical protein"
                     /gene="BMETHMGA300037"
     CDS             complement(48327..48959)
                     /note="hypothetical protein"
                     /gene="BMETHMGA300038c"
     CDS             49206..50108
                     /note="hypothetical protein"
                     /gene="BMETHMGA300039"
     CDS             50434..51555
                     /note="GTP-dependent nucleic acid binding protein"
                     /gene="BMETHMGA300040"
     CDS             51782..52300
                     /note="phosphoesterase PA-phosphoesterase-like
            protein"
                     /gene="BMETHMGA300041"
     CDS             52317..53471
                     /note="glycosyl transferase, group I"
                     /gene="BMETHMGA300042"
     CDS             53638..53925
                     /note="30S ribosomal protein S6"
                     /gene="BMETHMGA300043"
     CDS             54002..54496
                     /note="Single-stranded DNA binding protein"
                     /gene="BMETHMGA300044"
     CDS             55438..56382
                     /note="hypothetical protein"
                     /gene="BMETHMGA300045"
```

```
CDS             56414..58387
                /note="phosphodiesterase"
                /gene="BMETHMGA300046"
CDS             58384..58833
                /note="50S ribosomal protein L9"
                /gene="BMETHMGA300047"
CDS             59026..60390
                /note="replicative DNA helicase"
                /gene="BMETHMGA300048"
CDS             60630..61952
                /note="adenylosuccinate synthetase"
                /gene="BMETHMGA300049"
CDS             62221..63771
                /note="multidrug resistance protein"
                /gene="BMETHMGA300050"
fasta_record    64148..73513
                /note="ConsensusfromContig5 Average coverage: 527.34"
                /label=ConsensusfromContig5 Average coverage: 527.34
                /colour=10
CDS             64929..67298
                /note="Lysylcardiolipin synthase/lysyltransferase"
                /gene="BMETHMGA300051"
CDS             complement(67520..68320)
                /note="modulator of CcpN activity"
                /gene="BMETHMGA300052c"
CDS             complement(68367..68858)
                /note="transcriptional repressor CcpN"
                /gene="BMETHMGA300053c"
CDS             complement(69210..70592)
                /note="glycyl=tRNA synthetase"
                /gene="BMETHMGA300054c"
CDS             complement(70884..71639)
                /note="DNA repair protein RecO"
                /gene="BMETHMGA300055c"
CDS             complement(71974..72777)
                /note="GTP-binding protein Era"
                /gene="BMETHMGA300056c"
CDS             complement(72887..73279)
                /note="cytidine deaminase"
                /gene="BMETHMGA300057c"
fasta_record    73514..74844
                /note="ConsensusfromContig6 Average coverage: 479.39"
                /label=ConsensusfromContig6 Average coverage: 479.39
                /colour=11
fasta_record    74845..76134
                /note="ConsensusfromContig7 Average coverage: 579.55"
                /label=ConsensusfromContig7 Average coverage: 579.55
                /colour=10
CDS             complement(75234..75515)
                /note="hypothetical protein"
                /gene="BMETHMGA300058c"
fasta_record    76135..78201
                /note="ConsensusfromContig8 Average coverage: 484.90"
```

```
                /label=ConsensusfromContig8 Average coverage: 484.90
                /colour=11
CDS             76591..77034
                /note="hypothetical protein"
                /gene="BMETHMGA300059"
CDS             77354..77695
                /note="hypothetical protein"
                /gene="BMETHMGA300060"
fasta_record    78202..79577
                /note="ConsensusfromContig9 Average coverage: 926.99"
                /label=ConsensusfromContig9 Average coverage: 926.99
                /colour=10
CDS             complement(78611..79156)
                /note="hypothetical protein"
                /gene="BMETHMGA300062c"
fasta_record    79578..80106
                /note="ConsensusfromContig10 Average coverage:
       1,078.66"
                /label=ConsensusfromContig10 Average coverage:
       1,078.66
                /colour=11
CDS             79673..80209
                /note="cobalamin biosynthesis CbiX"
                /gene="BMETHMGA300063"
fasta_record    80107..81098
                /note="ConsensusfromContig11 Average coverage:
       511.04"
                /label=ConsensusfromContig11 Average coverage: 511.04
                /colour=10
CDS             80252..80899
                /note="transposase mutator type"
                /gene="BMETHMGA300064"
fasta_record    81099..260066
                /note="ConsensusfromContig12 Average coverage:
       504.69"
                /label=ConsensusfromContig12 Average coverage: 504.69
                /colour=11
CDS             81422..82933
                /note="thermostable carboxypeptidase I"
                /gene="BMETHMGA300065"
CDS             83050..83991
                /note="Hypothetical protein"
                /gene="BMETHMGA300066"
CDS             complement(84037..85692)
                /note="cytochrome c oxidase subunit I"
                /gene="BMETHMGA300067c"
CDS             complement(85692..86168)
                /note="cytochrome c oxidase subunit II"
                /gene="BMETHMGA300068c"
CDS             86850..87587
                /note="xanthine phosphoribosyltransferase"
                /gene="BMETHMGA300069"
CDS             87584..88897
```

```
                     /note="xanthine permease"
                     /gene="BMETHMGA300070"
     CDS             complement(89099..89743)
                     /note="alkaline phosphatase"
                     /gene="BMETHMGA300071c"
     CDS             90487..91572
                     /note="naringenin-chalcone synthase"
                     /gene="BMETHMGA300072"
     CDS             91572..92135
                     /note="hypothetical protein"
                     /gene="BMETHMGA300073"
     CDS             92491..96156
                     /note="reticuloyte binding protein"
                     /gene="BMETHMGA300074"
     CDS             96170..96937
                     /note="hypothetical protein"
                     /gene="BMETHMGA300075"
     CDS             98076..99245
                     /note="cell division protein FtsZ"
                     /gene="BMETHMGA300077"
     CDS             99461..100387
                     /note="5'-3' exonuclease"
                     /gene="BMETHMGA300078"
     CDS             complement(101527..102450)
                     /note="transporter EamA"
                     /gene="BMETHMGA300079c"
     CDS             complement(102470..102883)
                     /note="ribonuclease H"
                     /gene="BMETHMGA300080c"
     CDS             102999..103655
                     /note="ribonuclease H"
                     /gene="BMETHMGA300081"
     CDS             104384..104674
                     /note="hypothetical protein"
                     /gene="BMETHMGA300082"
     CDS             104765..105283
                     /note="hypothetical protein"
                     /gene="BMETHMGA300083"
     CDS             105428..106795
                     /note="glycolate oxidase subunit GlcD"
                     /gene="BMETHMGA300084"
     CDS             106962..109145
                     /note="DNA topoisomerase III"
                     /gene="BMETHMGA300085"
     CDS             109672..110388
                     /note="Hypothetical protein"
                     /gene="BMETHMGA300086"
     CDS             110475..110909
                     /note="manganase transport transcriptional regulator"
                     /gene="BMETHMGA300087"
     CDS             111288..111893
                     /note="hypothetical protein"
                     /gene="BMETHMGA300088"
```

```
CDS             complement(112029..113156)
                /note="diacylglycerol glucosyltransferase"
                /gene="BMETHMGA300089c"
CDS             complement(113325..114167)
                /note="homoserine O-succinyltransferase"
                /gene="BMETHMGA300090c"
CDS             114414..116102
                /note="formate-tetrahydrofolate ligase"
                /gene="BMETHMGA300091"
CDS             116141..116674
                /note="hypothetical protein"
                /gene="BMETHMGA300092"
CDS             117025..117876
                /note="DegV family protein"
                /gene="BMETHMGA300093"
CDS             117939..118604
                /note="metal-dependent phoshpohydrolase"
                /gene="BMETHMGA300094"
CDS             118586..120484
                /note="ABC transporter ATP-binding protein"
                /gene="BMETHMGA300095"
CDS             120635..121789
                /note="Putative lyase YpgR"
                /gene="BMETHMGA300096"
CDS             122053..123765
                /note="dihydroxy-acid dehydratase"
                /gene="BMETHMGA300097"
CDS             123843..124277
                /note="hypothetical protein"
                /gene="BMETHMGA300098"
CDS             124408..125184
                /note="protein-L-isoD(D-D) O-methyltransferase"
                /gene="BMETHMGA300099"
CDS             125590..126213
                /note="hypothetical protein"
                /gene="BMETHMGA300100"
CDS             126405..127178
                /note="integral membrane protein TerC"
                /gene="BMETHMGA300101"
CDS             127318..128028
                /note="anthrax toxin A moiety lethal factor"
                /gene="BMETHMGA300102"
CDS             128269..128964
                /note="thymidylate synthase"
                /gene="BMETHMGA300103"
CDS             128961..129446
                /note="dihyrofolate reductase"
                /gene="BMETHMGA300104"
CDS             129663..130931
                /note="threonine dehydratase"
                /gene="BMETHMGA300105"
CDS             131405..132073
                /note="Sec-independent protein translocase, TatC"
```

```
                    /gene="BMETHMGA300106"
     CDS            132707..133285
                    /note="cytochrome c oxidase Cu(A) center assembly
         protein"
                    /gene="BMETHMGA300107"
     CDS            133333..134229
                    /note="putative exported lipase/acylhydrolase"
                    /gene="BMETHMGA300108"
     CDS            134226..134804
                    /note="hypothetical protein"
                    /gene="BMETHMGA300109"
     CDS            complement(134908..136311)
                    /note="RNA methylase"
                    /gene="BMETHMGA300110c"
     CDS            136445..137473
                    /note="peptide methionine sulfoxide reductase"
                    /gene="BMETHMGA300111"
     CDS            complement(137518..138114)
                    /note="hypothetical protein"
                    /gene="BMETHMGA300112c"
     CDS            complement(138263..139348)
                    /note="hypothetical protein"
                    /gene="BMETHMGA300113c"
     CDS            140043..141431
                    /note="FAD dependent oxidoreducatase"
                    /gene="BMETHMGA300114"
     CDS            142376..143731
                    /note="aminotransferase, class III"
                    /gene="BMETHMGA300116"
     CDS            143732..144424
                    /note="acetate CoA-transferase"
                    /gene="BMETHMGA300117"
     CDS            144409..145071
                    /note="3-oxoacid CoA-transferase, B subunit"
                    /gene="BMETHMGA300118"
     CDS            145104..146360
                    /note="acetylornithine deacetylase"
/gene="BMETHMGA300119"
     CDS            146437..147291
                    /note="Beta-lysine acetyltransferase"
/gene="BMETHMGA300120"
     CDS            147322..148770
                    /note="arginine utilization regulatory protein"
                    /gene="BMETHMGA300121"
     CDS            148920..150371
                    /note="L-lysine 2,3-aminomutase"
                    /gene="BMETHMGA300122"
     CDS            150408..150692
                    /note="hypothetical protein"
                    /gene="BMETHMGA300124"
     CDS            151261..152049
                    /note="hypothetical protein"
                    /gene="BMETHMGA300125"
```

```
CDS             152822..153514
                /note="sporulation control protein"
                /gene="BMETHMGA300126"
CDS             153719..154027
                /note="hypothetical protein"
                /gene="BMETHMGA300127"
CDS             154561..155268
                /note="purine-nucleoside phosphorylase"
                /gene="BMETHMGA300128"
CDS             155475..156308
                /note="carboxypeptidase"
                /gene="BMETHMGA300129"
CDS             complement(156465..157247)
                /note="bacitracin resistance protein BacA"
                /gene="BMETHMGA300130c"
CDS             157411..157743
                /note="hypothetical protein"
                /gene="BMETHMGA300131"
CDS             158166..158894
                /note="hypothetical protein"
                /gene="BMETHMGA300132"
CDS             158907..160616
                /note="glucose-methanol-choline oxidoreductase"
                /gene="BMETHMGA300133"
CDS             160926..161732
                /note="hypothetical protein"
                /gene="BMETHMGA300134"
CDS             161892..162467
                /note="thiol:disulfide interchange protein
       (thioredoxin)"
                /gene="BMETHMGA300135"
CDS             complement(162886..163371)
                /note="putative membrane protein"
                /gene="BMETHMGA300136c"
CDS             complement(164224..167316)
                /note="acriflavin resistance protein"
                /gene="BMETHMGA300137c"
CDS             complement(167456..170206)
                /note="YVTN beta-propeller repeat-containing protein"
                /gene="BMETHMGA300138c"
CDS             170969..172084
                /note="riboflavin biosynthesis protein RibD"
                /gene="BMETHMGA300139"
CDS             172069..172719
                /note="riboflavin synthase subunit alpha"
                /gene="BMETHMGA300140"
CDS             172720..173931
                /note="bifunctional 3,4-dihydroxy-2-butanone
                4-synthase/GTP cyclohydrolase II protein "
                /gene="BMETHMGA300141"
CDS             173953..174420
                /note="riboflavin synthase subunit beta"
                /gene="BMETHMGA300142"
```

```
CDS             174540..175355
                /note="hypothetical protein"
                /gene="BMETHMGA300143"
CDS             175598..176392
                /note="diaminopimelate epimerase"
                /gene="BMETHMGA300144"
CDS             complement(176482..178311)
                /note="DNA ligase D"
                /gene="BMETHMGA300145c"
CDS             complement(178558..178893)
                /note="ErfK/YbiS/YcfS/YnhG family protein"
                /gene="BMETHMGA300146c"
CDS             complement(179588..180205)
                /note="Alkaline phosphatase, DedA"
                /gene="BMETHMGA300147c"
CDS             complement(180676..181083)
                /note="AIG2 family protein"
                /gene="BMETHMGA300148c"
CDS             181868..182104
                /note="hypothetical protein"
                /gene="BMETHMGA300149"
CDS             182350..182655
                /note="hypothetical protein"
                /gene="BMETHMGA300150"
CDS             182843..183775
                /note="Mg2 transporter protein CorA family protein"
                /gene="BMETHMGA300151"
CDS             184317..185153
                /note="ATP-dependent DNA ligase"
                /gene="BMETHMGA300152"
CDS             185251..186981
                /note="glycerol-3-phosphate dehydrogenase"
                /gene="BMETHMGA300153"
CDS             187869..188177
                /note="hypothetical protein"
                /gene="BMETHMGA300154"
CDS             complement(188302..188592)
                /note="hypothetical protein"
                /gene="BMETHMGA300155c"
CDS             188782..190050
                /note="signal transduction histidine kinase"
                /gene="BMETHMGA300156"
CDS             complement(190327..191082)
                /note="hypothetical protein"
                /gene="BMETHMGA300157c"
CDS             191298..192968
                /note="sulfate transporter"
                /gene="BMETHMGA300158"
CDS             192982..193335
                /note="ArsR family transcriptional regulator"
                /gene="BMETHMGA300159"
CDS             193898..194911
                /note="major facilitator superfamily protein"
```

```
                     /gene="BMETHMGA300160"
CDS                  195921..196697
                     /note="hypothetical protein"
                     /gene="BMETHMGA300161"
CDS                  196901..197434
                     /note="Activator of Hsp90 ATPase I family protein"
                     /gene="BMETHMGA300162"
CDS                  197966..198340
                     /note="carbohydrate-binding and sugar hydrolysis"
                     /gene="BMETHMGA300163"
CDS                  198780..200174
                     /note="iron-regulated membrane protein"
                     /gene="BMETHMGA300164"
CDS                  200574..201284
                     /note="Sec-independent protein translocase"
                     /gene="BMETHMGA300165"
CDS                  201594..201977
                     /note="hypothetical protein"
                     /gene="BMETHMGA300166"
CDS                  complement(202233..203684)
                     /note="GntR family transcriptional regulator"
                     /gene="BMETHMGA300167c"
CDS                  203869..204495
                     /note="amino acid transporter LysE"
                     /gene="BMETHMGA300168"
CDS                  204835..205212
                     /note="N-acetyltransferase GCN5"
                     /gene="BMETHMGA300169"
CDS                  206074..207984
                     /note="hypothetical protein"
                     /gene="BMETHMGA300171"
CDS                  208069..208503
                     /note="BadM/RrF2 family transcriptional regulator"
                     /gene="BMETHMGA300172"
CDS                  208728..209129
                     /note="DoxX family protein"
                     /gene="BMETHMGA300173"
CDS                  209692..210240
                     /note="DeoR family transcriptional regulator"
                     /gene="BMETHMGA300174"
CDS                  210237..211142
                     /note="1-phosphofructokinase"
                     /gene="BMETHMGA300175"
CDS                  211462..213030
                     /note="PTS system, fructose specific transporter"
                     /gene="BMETHMGA300176"
CDS                  213583..214380
                     /note="hypothetical protein"
                     /gene="BMETHMGA300177"
CDS                  215061..215855
                     /note="putative hydrolase"
                     /gene="BMETHMGA300178"
CDS                  complement(216052..217317)
```

```
                     /note="multidrug resistance protein"
                     /gene="BMETHMGA300179c"
     CDS             217566..218057
                     /note="MOSC domain-containing protein"
                     /gene="BMETHMGA300180"
     CDS             219179..220564
                     /note="anthranilate synthase component I"
                     /gene="BMETHMGA300181"
     CDS             220561..221172
                     /note="anthranilate synthase component II"
                     /gene="BMETHMGA300182"
     CDS             221144..222169
                     /note="antharnilate phosphoribosyltransferase"
                     /gene="BMETHMGA300183"
     CDS             222170..222958
                     /note="indole-3-gylcerol phosphate synthase"
                     /gene="BMETHMGA300184"
     CDS             222955..223566
                     /note="N-(5'-phosphoribosyl) anthranilate isomerase"
                     /gene="BMETHMGA300185"
     CDS             223563..224768
                     /note="tryptophan synthase, subunit beta"
                     /gene="BMETHMGA300186"
     CDS             224765..225547
                     /note="tryptohan synthase subunit alpha"
                     /gene="BMETHMGA300187"
     CDS             228039..228617
                     /note="hypothetical protein"
                     /gene="BMETHMGA300188"
     CDS             complement(228768..229625)
                     /note="DNA repair exonuclease"
                     /gene="BMETHMGA300189c"
     CDS             229768..230889
                     /note="monogalactosyldiacylglycerol synthase"
                     /gene="BMETHMGA300190"
     CDS             231228..233546
                     /note="aldehyde oxidase and xanthine
            dehydrogenase/carbon
                     monoxide dehydrogenase large subunit"
                     /gene="BMETHMGA300191"
     CDS             233689..234381
                     /note="xanthine dehydrogenase FAD-binding
            subunit/carbon
                     monoxide dehydrogenase middle subunit"
                     /gene="BMETHMGA300192"
     CDS             234382..234876
                     /note="carbon monoxide dehydrogenase small subunit"
                     /gene="BMETHMGA300193"
     CDS             235213..236358
                     /note="glycerate kinase"
                     /gene="BMETHMGA300194"
     CDS             236528..236869
                     /note="hypothetical protein"
```

```
                     /gene="BMETHMGA300195"
     CDS             237449..239179
                     /note="lipid A export ATP-binding/permease protein"
                     /gene="BMETHMGA300196"
     CDS             239226..240110
                     /note="hypothetical protein"
                     /gene="BMETHMGA300197"
     CDS             240214..241713
                     /note="PAS domain S-box protein"
                     /gene="BMETHMGA300198"
     CDS             241726..242373
                     /note="two component LuxR family transcriptional
                     regulator"
                     /gene="BMETHMGA300199"
     CDS             242379..242708
                     /note="hypothetical protein"
                     /gene="BMETHMGA300200"
     CDS             complement(242925..243527)
                     /note="two component response regulator"
                     /gene="BMETHMGA300201c"
     CDS             complement(243524..244675)
                     /note="two-component signal transduction histidine
            kinase"
                     /gene="BMETHMGA300202c"
     CDS             complement(244683..245420)
                     /note="ABC transporter permease"
                     /gene="BMETHMGA300203c"
     CDS             complement(245420..246325)
                     /note="ABC transporter ATP binding subunit"
                     /gene="BMETHMGA300204c"
     CDS             246747..247757
                     /note="ABC transporter, substrate binding subunit"
                     /gene="BMETHMGA300205"
     CDS             247770..248063
                     /note="hypothetical protein"
                     /gene="BMETHMGA300206"
     CDS             248060..248827
                     /note="nitrate/sulfonate/bicarbonate ABC transporter
                     permease"
                     /gene="BMETHMGA300207"
     CDS             248824..249564
                     /note="putative nitrate/sulfonate/biocarbonate ABC
                     transporter"
                     /gene="BMETHMGA300208"
     CDS             249519..250337
                     /note="TatD related dioxyribonuclease"
                     /gene="BMETHMGA300209"
     CDS             250431..251198
                     /note="phosphatase"
                     /gene="BMETHMGA300211"
     CDS             252011..253054
                     /note="putative dehydrogenase"
                     /gene="BMETHMGA300212"
```

```
     CDS             253231..254859
                     /note="major facilitator family transporter"
                     /gene="BMETHMGA300213"
     CDS             255061..255978
                     /note="hypothetical protein"
                     /gene="BMETHMGA300214"
     CDS             complement(256104..256406)
                     /note="transition state regulator"
                     /gene="BMETHMGA300215c"
     CDS             256799..257254
                     /note="PTS system EIIA component"
                     /gene="BMETHMGA300216"
     CDS             257275..257577
                     /note="PTS system mannitol specific enzyme II"
                     /gene="BMETHMGA300217"
     CDS             257601..258971
                     /note="PTS system ascorbate-specific permease"
/gene="BMETHMGA300218"
     CDS             259056..259592
                     /note="KR domain protein"
                     /gene="BMETHMGA300219"
     CDS             259537..259986
                     /note="KR domain protein"
/gene="BMETHMGA300220"
     CDS             complement(260057..260320)
                     /note="chorismate synthase"
                     /gene="BMETHMGA300221c"
     fasta_record    260067..270440
                     /note="ConsensusfromContig13 Average coverage:
            518.29"
                     /label=ConsensusfromContig13 Average coverage: 518.29
                     /colour=10
     CDS             complement(260295..260543)
                     /note="chorismate synthase"
                     /gene="BMETHMGA300222c"
     CDS             complement(260949..261428)
                     /note="3-dehydroquinate synthase"
                     /gene="BMETHMGA300223c"
     CDS             complement(261731..262921)
                     /note="chorismate synthase"
                     /gene="BMETHMGA300224c"
     CDS             264245..265330
                     /note="nucleoside permease NupC"
                     /gene="BMETHMGA300225"
     CDS             complement(265564..266910)
                     /note="sulfite reductase subunit beta"
                     /gene="BMETHMGA300226c"
     CDS             complement(266838..267281)
                     /note="sulfite reductase subunit beta"
                     /gene="BMETHMGA300228c"
     CDS             complement(267308..269194)
                     /note="sufite reductase flavoprotein alpha subunit"
                     /gene="BMETHMGA300229c"
```

```
CDS             269480..270310
                /note="Regulator of sulfur assimilation CysL"
                /gene="BMETHMGA300230"
fasta_record    270441..295687
                /note="ConsensusfromContig14 Average coverage:
       616.35"
                /label=ConsensusfromContig14 Average coverage: 616.35
                /colour=11
CDS             270698..271777
                /note="glycosyl transferase family protein"
                /gene="BMETHMGA300231"
CDS             272055..273167
                /note="gluatmine-scyllo-inositol transaminase"
                /gene="BMETHMGA300232"
CDS             273237..273986
                /note="lipopolysaccharide biosynthesis protein"
                /gene="BMETHMGA300233"
CDS             273958..274992
                /note="lipopolysaccharide biosynthesis protein"
                /gene="BMETHMGA300234"
CDS             275006..276319
                /note="UDP-N-acetyl-D-mannosaminuuronate
       dehydrogenase"
                /gene="BMETHMGA300235"
CDS             276347..277078
                /note="acetyltransferase"
                /gene="BMETHMGA300236"
CDS             277131..278510
                /note="polysaccharide biosynthesis protein"
                /gene="BMETHMGA300237"
CDS             278535..279593
                /note="UDP-N-acetylglucosamine 2-epimerase"
                /gene="BMETHMGA300238"
CDS             279608..280996
                /note="hypothetical protein"
                /gene="BMETHMGA300239"
CDS             281000..282112
                /note="hypothetical protein/glycosyl transferase
       group I"
                /gene="BMETHMGA300240"
CDS             282109..283191
                /note="glycosyltransferase"
                /gene="BMETHMGA300241"
CDS             283188..284051
                /note="UTP-glucose-1-phosphate uridyltransferase"
                /gene="BMETHMGA300242"
CDS             284065..285300
                /note="glycosyl transferase group I"
                /gene="BMETHMGA300243"
CDS             complement(285408..286721)
                /note="hypothetical protein"
                /gene="BMETHMGA300244c"
CDS             complement(286723..287418)
```

```
                /note="hypothetical protein"
                /gene="BMETHMGA300245c"
CDS             complement(287547..289886)
                /note="N-acetylmuramoyl-L-alanine amidase"
                /gene="BMETHMGA300246c"
CDS             complement(290130..291158)
                /note="LytR gene product"
                /gene="BMETHMGA300247c"
CDS             complement(291263..291898)
                /note="hypothetical protein"
                /gene="BMETHMGA300248c"
CDS             292287..293426
                /note="two component sensor histidine kinase"
                /gene="BMETHMGA300249"
CDS             293625..294149
                /note="two-component response regulator"
                /gene="BMETHMGA300250"
CDS             294443..295285
                /note="degV family protein"
                /gene="BMETHMGA300251"
CDS             295344..295694
                /note="hypothetical protein"
                /gene="BMETHMGA300252"
fasta_record    295688..413509
                /note="ConsensusfromContig15 Average coverage:
       550.00"
                /label=ConsensusfromContig15 Average coverage: 550.00
                /colour=10
CDS             complement(296272..296784)
                /note="hypothetical protein"
                /gene="BMETHMGA300253c"
CDS             complement(296844..297548)
                /note="hypothetical protein"
                /gene="BMETHMGA300254c"
CDS             complement(298032..298466)
                /note="transposase"
                /gene="BMETHMGA300255c"
CDS             complement(298527..298958)
                /note="transposase mutator type"
                /gene="BMETHMGA300256c"
CDS             complement(298940..299275)
                /note="transposase mutator type"
                /gene="BMETHMGA300257c"
CDS             complement(299831..300400)
                /note="phosphopantetheinyl transferase"
                /gene="BMETHMGA300258c"
CDS             complement(300419..303367)
                /note="non-ribosomal peptide synthetase module"
                /gene="BMETHMGA300259c"
CDS             complement(303426..305159)
                /note="methyl-accepting chemotaxis protein"
                /gene="BMETHMGA300260c"
CDS             306679..307329
```

```
                     /note="integrase"
                     /gene="BMETHMGA300261"
     CDS             307509..307832
                     /note="phage integrase"
                     /gene="BMETHMGA300262"
     CDS             complement(308046..309464)
                     /note="site-specific recombinase"
/gene="BMETHMGA300263c"
     CDS             310511..311239
                     /note="glutamine ABC transporter ATP-binding protein"
                     /gene="BMETHMGA300264"
     CDS             311257..312090
                     /note="glutamine ABC transporter glutamine-binding
                     protein"
                     /gene="BMETHMGA300265"
     CDS             312246..312899
                     /note="glutamine ABC transporter permease"
                     /gene="BMETHMGA300266"
     CDS             312913..313563
                     /note="glutamine ABC transporter membrane protein"
                     /gene="BMETHMGA300267"
CDS             313864..314379
                     /note="RNA ligase or phosphoesterase"
                     /gene="BMETHMGA300268"
     CDS             314419..314805
                     /note="putative acetyltransferase"
                     /gene="BMETHMGA300269"
     CDS             complement(315357..315614)
                     /note="spoVIF"
                     /gene="BMETHMGA300270c"
     CDS             complement(316163..316519)
                     /note="stage V sporulation protein AE"
                     /gene="BMETHMGA300271c"
     CDS             complement(316535..317554)
                     /note="stage V spurulation protein AD"
                     /gene="BMETHMGA300272c"
     CDS             complement(317554..318033)
                     /note="stage V sporulation protein AC"
                     /gene="BMETHMGA300273c"
     CDS             complement(318049..318537)
                     /note="sporulation lipoprotein"
                     /gene="BMETHMGA300274c"
     CDS             complement(318774..319136)
                     /note="sporulation specific protein"
                     /gene="BMETHMGA300275c"
     CDS             complement(319860..320633)
                     /note="enoyl-(acyl carrier protein) reductase"
                     /gene="BMETHMGA300276c"
     CDS             320845..321585
                     /note="serine/threonine protein phosphatase"
                     /gene="BMETHMGA300277"
     CDS             complement(321631..322551)
                     /note="pseudouridine synthase"
```

```
                     /gene="BMETHMGA300278c"
     CDS             complement(322571..323371)
                     /note="inorganic polyphosphate/ATP-NAD kinase"
                     /gene="BMETHMGA300279c"
     CDS             complement(323417..323980)
                     /note="GTP pyrophosphokinase"
                     /gene="BMETHMGA300280c"
     CDS             complement(324121..324498)
                     /note="putative phosphatase"
                     /gene="BMETHMGA300281c"
     CDS             324617..325195
                     /note="adenylate cyclase"
                     /gene="BMETHMGA300282"
     CDS             325218..325847
                     /note="lytic murein transglycosylase"
                     /gene="BMETHMGA300283"
     CDS             326567..327367
                     /note="thioredoxin-like protein"
                     /gene="BMETHMGA300284"
     CDS             complement(327955..329778)
                     /note="oligoendopeptidase F"
                     /gene="BMETHMGA300285c"
     CDS             complement(329899..331095)
                     /note="competence CoiA family protein"
                     /gene="BMETHMGA300286c"
     CDS             complement(331195..332709)
                     /note="phospholipase D/transphosphatidylase"
                     /gene="BMETHMGA300287c"
     CDS             complement(332791..333459)
                     /note="adaptor protein"
                     /gene="BMETHMGA300288c"
     CDS             complement(333787..334182)
                     /note="transcriptional regulatory protein Spx"
                     /gene="BMETHMGA300289c"
     CDS             complement(335338..336273)
                     /note="oligopeptide ABC transporter"
                     /gene="BMETHMGA300290c"
     CDS             complement(336275..337342)
                     /note="oligopeptide ABC transporter"
                     /gene="BMETHMGA300291c"
     CDS             complement(337355..338272)
                     /note="oligopeptide ABC transporter permease"
                     /gene="BMETHMGA300292c"
     CDS             complement(338275..339225)
                     /note="oligopeptide ABC transporter permease"
                     /gene="BMETHMGA300293c"
     CDS             complement(339316..340929)
                     /note="ABC transporter periplasmic protein"
/gene="BMETHMGA300294c"
     CDS             341982..342971
                     /note="tryptophanyl-tRNA synthetase"
/gene="BMETHMGA300295"
     CDS             complement(343005..343757)
```

```
                     /note="hypothetical protein"
                     /gene="BMETHMGA300296c"
     CDS             complement(344226..344879)
                     /note="hypothetical protein"
                     /gene="BMETHMGA300297c"
     CDS             complement(345135..346376)
                     /note="3-oxoacyl-(acyl carrier protein) synthase"
                     /gene="BMETHMGA300298c"
     CDS             complement(346648..347580)
                     /note="3-oxoacyl-(acyl carrier protein) synthase"
                     /gene="BMETHMGA300299c"
     CDS             347797..348099
                     /note="3-oxoacyl synthase"
                     /gene="BMETHMGA300300"
     CDS             complement(348158..349117)
                     /note="transcriptional activator protein Med"
                     /gene="BMETHMGA300301c"
     CDS             complement(349276..350139)
                     /note="glucose epimerase"
                     /gene="BMETHMGA300302c"
     CDS             complement(350220..350987)
                     /note="hypothetical protein"
                     /gene="BMETHMGA300303c"
     CDS             complement(351476..354115)
                     /note="ATP-dependent chaperone ClpB"
                     /gene="BMETHMGA300304c"
     CDS             complement(354523..355512)
                     /note="ornithine carbamoyltransferase"
                     /gene="BMETHMGA300306c"
     CDS             complement(355473..358586)
                     /note="carbamoyl-phosphate synthase, large subunit"
                     /gene="BMETHMGA300307c"
     CDS             complement(358579..359661)
                     /note="carbamoyl-phosphate synthase, small subunit"
                     /gene="BMETHMGA300308c"
     CDS             complement(359729..360892)
                     /note="acetylornithine and succinylornithine
                     aminotransferase"
                     /gene="BMETHMGA300309c"
     CDS             complement(360889..361662)
                     /note="acetylglutamate kinase"
                     /gene="BMETHMGA300310c"
     CDS             complement(361675..362976)
                     /note="bifunctional ornithine
                     acetyltransferase/N-acetylglutamate synthase"
                     /gene="BMETHMGA300311c"
     CDS             complement(362927..363961)
                     /note="N-acetyl-gamma-glutamyl-phosphate reductase"
                     /gene="BMETHMGA300312c"
     CDS             complement(364127..364681)
                     /note="molybdopterin-guanine dinucleotide
          biosynthesis
                     protein B"
```

```
                     /gene="BMETHMGA300313c"
CDS                  complement(364669..365940)
                     /note="molybdenum cofactor synthesis domain protein"
                     /gene="BMETHMGA300314c"
CDS                  complement(366007..367038)
                     /note="molybdenum cofactor biosynthesis protein A"
                     /gene="BMETHMGA300315c"
CDS                  complement(367302..367610)
                     /note="hypothetical protein"
                     /gene="BMETHMGA300316c"
CDS                  complement(367633..368421)
                     /note="putative hydrolase"
                     /gene="BMETHMGA300317c"
CDS                  368626..369444
                     /note="putative phosphatase YitU"
                     /gene="BMETHMGA300318"
CDS                  complement(369958..370308)
                     /note="putative intracellular proteinase inhibitor"
                     /gene="BMETHMGA300319c"
CDS                  complement(370387..371229)
                     /note="putative integral membrane protein YitT"
                     /gene="BMETHMGA300320c"
CDS                  371407..372264
                     /note="YitS protein"
                     /gene="BMETHMGA300321"
CDS                  complement(372636..373577)
                     /note="RNA binding S1 domain protein"
                     /gene="BMETHMGA300322c"
CDS                  373629..374138
                     /note="nucleotide-binding protein"
                     /gene="BMETHMGA300323"
CDS                  complement(374275..376635)
                     /note="alpha-glucosidase"
                     /gene="BMETHMGA300324c"
CDS                  complement(376706..378241)
                     /note="alpha-amylase"
                     /gene="BMETHMGA300325c"
CDS                  378428..380221
                     /note="alpha amylase catalytic subunit"
                     /gene="BMETHMGA300326"
CDS                  complement(380298..382154)
                     /note="asparagine synthase"
                     /gene="BMETHMGA300327c"
CDS                  382178..382876
                     /note="Hypothetical protein"
                     /gene="BMETHMGA300328"
CDS                  complement(382991..384082)
                     /note="ABC transporter"
                     /gene="BMETHMGA300329c"
CDS                  complement(384079..385347)
                     /note="ABC-2 type transporter"
                     /gene="BMETHMGA300330c"
CDS                  complement(385359..386291)
```

```
/note="ABC transporter"
                     /gene="BMETHMGA300331c"
CDS              complement(386407..387039)
                     /note="two-component response regulator, LuxR family"
                     /gene="BMETHMGA300332c"
     CDS             complement(387017..388159)
                     /note="two=-component sensor histidine kinase"
                     /gene="BMETHMGA300333c"
     CDS             complement(388325..388705)
                     /note="hypothetical protein"
                     /gene="BMETHMGA300334c"
     CDS             complement(388768..389679)
                     /note="5-carboxymethyl-2-hydroxymuconate delta-
            isomerase"
                     /gene="BMETHMGA300335c"
     CDS             complement(389721..390914)
                     /note="hypothetical protein"
                     /gene="BMETHMGA300336c"
     CDS             391877..392452
                     /note="spore germination protein GerPC"
                     /gene="BMETHMGA300337"
     CDS             392621..393022
                     /note="germination protein"
                     /gene="BMETHMGA300338"
     CDS             393079..393309
                     /note="germination protein PF"
                     /gene="BMETHMGA300339"
     CDS             complement(393352..397134)
                     /note="recombination helicase AddA"
                     /gene="BMETHMGA300340c"
     CDS             complement(397094..400615)
                     /note="ATP-dependent nuclease subunit B"
                     /gene="BMETHMGA300341c"
     CDS             complement(400899..401453)
                     /note="signal peptidase I"
                     /gene="BMETHMGA300342c"
     CDS             complement(401491..402102)
                     /note="hypothetical protein"
                     /gene="BMETHMGA300343c"
     CDS             complement(402213..403187)
                     /note="transcriptional regulatory protein MsrR"
                     /gene="BMETHMGA300344c"
     CDS             403479..403919
                     /note="hypothetical protein"
                     /gene="BMETHMGA300345"
     CDS             complement(403951..404529)
                     /note="competence transcription factor"
                     /gene="BMETHMGA300346c"
     CDS             complement(405146..406279)
                     /note="peptidase M48 "
                     /gene="BMETHMGA300347c"
     CDS             complement(406652..406936)
                     /note="branched chain amino acid transport"
```

```
                     /gene="BMETHMGA300348c"
     CDS             complement(406947..407678)
                     /note="branched chain aminoacid transporter"
                     /gene="BMETHMGA300349c"
     CDS             407816..408367
                     /note="XRE family transcriptional regulator"
                     /gene="BMETHMGA300350"
     CDS             complement(408441..409262)
                     /note="ABC transporter related protein"
                     /gene="BMETHMGA300351c"
     CDS             complement(409210..410220)
/note="ABC transporter"
                     /gene="BMETHMGA300352c"
CDS             complement(410259..411260)
                     /note="ABC transporter"
                     /gene="BMETHMGA300353c"
     CDS             complement(411461..412252)
                     /note="enoxy-CoA hydratase family protein"
                     /gene="BMETHMGA300354c"
     CDS             complement(412579..413400)
                     /note="transposase, IS605 OrfB family"
                     /gene="BMETHMGA300355c"
     fasta_record    413510..469629
                     /note="ConsensusfromContig16 Average coverage:
            584.29"
                     /label=ConsensusfromContig16 Average coverage: 584.29
                     /colour=11
     CDS             413886..414230
                     /note="transcriptional protein"
                     /gene="BMETHMGA300356"
     CDS             414364..415335
                     /note="hypothetical protein"
                     /gene="BMETHMGA300357"
     CDS             complement(415492..415809)
                     /note="small acid-soluble protein alpha/beta"
                     /gene="BMETHMGA300358c"
     CDS             complement(415963..416325)
                     /note="glyoxalase family protein"
                     /gene="BMETHMGA300359c"
     CDS             416869..417780
                     /note="ABC transporter"
                     /gene="BMETHMGA300360"
     CDS             417777..418529
                     /note="putative permease"
                     /gene="BMETHMGA300361"
     CDS             418542..419267
                     /note="ABC transpoter permease"
                     /gene="BMETHMGA300362"
     CDS             419279..420001
                     /note="two-component response regulator"
                     /gene="BMETHMGA300363"
     CDS             420075..421433
```

```
                  /note="putative sensory transduction histidine
         kinase"
                  /gene="BMETHMGA300364"
     CDS          complement(421468..422445)
                  /note="Mg chelatase subunit ChlI"
                  /gene="BMETHMGA300365c"
     CDS          423414..425168
                  /note="signal transduction histidine kinase LytS"
                  /gene="BMETHMGA300366"
     CDS          425146..425892
                  /note="LytTR family two component trasnscriptional
                  regulator"
                  /gene="BMETHMGA300367"
     CDS          426085..427878
                  /note="carbon starvation protein CstA"
                  /gene="BMETHMGA300368"
     CDS          428066..429532
                  /note="amino acid transporter"
                  /gene="BMETHMGA300369"
     CDS          complement(429682..431043)
                  /note="hypothetical protein"
                  /gene="BMETHMGA300370c"
     CDS          complement(431061..431336)
                  /note="amino acid-binding ACT domain protein"
                  /gene="BMETHMGA300371c"
     CDS          432314..433129
                  /note="phosphomethylpyrimidine kinase"
                  /gene="BMETHMGA300372"
     CDS          433134..433805
                  /note="thiamine-phosphate pyrophosphorylase"
                  /gene="BMETHMGA300373"
     CDS          433879..435084
                  /note="ATP-dependent RNA helicase"
                  /gene="BMETHMGA300374"
     CDS          435789..436985
                  /note="Cys/Met metabolism pyridoxal-phosphate-
         dependent
                  protein"
                  /gene="BMETHMGA300375"
     CDS          437326..439125
                  /note="O-acetylhomoserine sulfhydrolase"
                  /gene="BMETHMGA300376"
     CDS          439157..440137
                  /note="nitrate/sulfonate/bicarbonate ABC transporter
periplasmic protein"
                  /gene="BMETHMGA300377"
     CDS          440193..441029
                  /note="binding-protein-dependent transporter innner
membrane component"
                  /gene="BMETHMGA300378"
     CDS          441034..441771
/note="ABC transporter ATP-binding protein"
/gene="BMETHMGA300379"
```

```
     CDS             441785..443497
                     /note="acetyl-CoA synthetase"
                     /gene="BMETHMGA300380"
     CDS             443635..444876
                     /note="homoserine dehydrogenase 2"
                     /gene="BMETHMGA300381"
     CDS             445317..445796
                     /note="S-adenosylmethionine:2-demethylmenaquinone
                     methyltransferase"
                     /gene="BMETHMGA300382"
     CDS             446013..446357
                     /note="hypothetical protein"
                     /gene="BMETHMGA300383"
     CDS             446377..447027
                     /note="phage shock protein PspA"
                     /gene="BMETHMGA300384"
     CDS             447116..447853
                     /note="hypothetical protein"
                     /gene="BMETHMGA300385"
     CDS             447850..448893
                     /note="two-component sensor histidine kinase"
                     /gene="BMETHMGA300386"
     CDS             448886..449518
                     /note="two-component response regulator"
                     /gene="BMETHMGA300387"
     CDS             complement(449629..451122)
                     /note="di-tripeptide ABC transporter"
                     /gene="BMETHMGA300388c"
     CDS             451366..452691
                     /note="polysaccharide biosynthesis protein"
                     /gene="BMETHMGA300389"
     CDS             453018..453515
                     /note="NADPH-dependent 7-cyano-7-deazoguanine
            reductase"
                     /gene="BMETHMGA300390"
     CDS             complement(453648..453923)
                     /note="acylphosphatase"
                     /gene="BMETHMGA300391c"
     CDS             complement(454097..455329)
                     /note="aminopeptidase"
                     /gene="BMETHMGA300392c"
     CDS             complement(456991..459084)
                     /note="DNA topoisomerase III"
                     /gene="BMETHMGA300394c"
     CDS             460146..460751
                     /note="hypothetical protein (phage DNA replication
protein)"
                     /gene="BMETHMGA300395"
     CDS             461029..461820
                     /note="phage protein"
/gene="BMETHMGA300396"
     CDS             462108..462365
                     /note="hypothetical protein"
```

```
                     /gene="BMETHMGA300397"
     CDS             complement(462457..462780)
                     /note="spore coat protein CotD"
                     /gene="BMETHMGA300398c"
     CDS             complement(464559..464915)
                     /note="hypothetical protein"
                     /gene="BMETHMGA300400c"
     CDS             466052..466387
                     /note="phage transcriptional regulator, ArpU"
                     /gene="BMETHMGA300402"
     CDS             466877..467191
                     /note="mannosyl-glycoprotein
                     endo-beta-N-acetylglucosamidase"
                     /gene="BMETHMGA300404"
     fasta_record    469630..544124
                     /note="ConsensusfromContig17 Average coverage:
            500.55"
                     /label=ConsensusfromContig17 Average coverage: 500.55
                     /colour=10
     CDS             470797..472464
                     /note="hypothetical protein"
                     /gene="BMETHMGA300407"
     CDS             complement(472907..475981)
                     /note="hypothtetical protein"
                     /gene="BMETHMGA300408c"
     CDS             complement(476947..477744)
                     /note="cell wall hydrolase phosphatase associtated
                     protein"
                     /gene="BMETHMGA300409c"
     CDS             complement(480186..480488)
                     /note="hypothetical protein"
                     /gene="BMETHMGA300411c"
     CDS             complement(480708..481346)
                     /note="TraR/DksA family transcriptional regulator"
                     /gene="BMETHMGA300412c"
     CDS             complement(481735..483090)
                     /note="aminotransferase class III"
                     /gene="BMETHMGA300413c"
     CDS             complement(483326..484366)
                     /note="alcohol dehydrogenase GroES, zinc-dependent"
                     /gene="BMETHMGA300414c"
     CDS             complement(484811..485356)
                     /note="acetyltransferase"
                     /gene="BMETHMGA300415c"
     CDS             complement(485690..487636)
                     /note="NADH:flavin oxidoreductase/NADH oxidase"
                     /gene="BMETHMGA300416c"
     CDS             complement(487893..489815)
                     /note="NAD:flavin oxidoreductase/NADH oxidase"
                     /gene="BMETHMGA300417c"
     CDS             complement(490129..490434)
                     /note="hypothetical protein"
                     /gene="BMETHMGA300418c"
```

```
CDS             complement(490520..490822)
                /note="hypothetical protein"
                /gene="BMETHMGA300419c"
CDS             491061..492053
                /note="quinone oxidoreductase, YhdH/YhfP family"
                /gene="BMETHMGA300420"
CDS             complement(492175..492531)
                /note="hypothetical protein"
                /gene="BMETHMGA300421c"
CDS             complement(492854..494020)
                /note="spore germination protein"
                /gene="BMETHMGA300422c"
CDS             complement(493945..494871)
                /note="spore germination protein"
                /gene="BMETHMGA300423c"
CDS             complement(495064..496557)
                /note="GerA spore germination protein"
                /gene="BMETHMGA300424c"
CDS             complement(496750..498195)
                /note="glucose-6-phosphate 1-dehydrogenase"
                /gene="BMETHMGA300425c"
CDS             498678..499313
                /note="3-hexulose-6-phosphate synthase"
                /gene="BMETHMGA300426"
CDS             499310..499873
                /note="6-phospho-3-hexuloisomerase"
                /gene="BMETHMGA300427"
CDS             complement(500099..501163)
                /note="coenzyme F420-dependent N5,N10-methylene
                teterahydromethanopterin reductase"
                /gene="BMETHMGA300428c"
CDS             complement(501287..502150)
                /note="triphenylmethane reductase"
                /gene="BMETHMGA300429c"
CDS             complement(502362..504260)
                /note="selenocysteine-specific translation elongation
                factor"
                /gene="BMETHMGA300430c"
CDS             complement(504290..505099)
                /note="pantothenate kinase"
                /gene="BMETHMGA300431c"
CDS             complement(505294..505728)
                /note="hypothetical protein"
                /gene="BMETHMGA300432c"
CDS             complement(505767..506447)
                /note="hypothetical protein"
                /gene="BMETHMGA300433c"
CDS             complement(506537..509674)
                /note="exonuclease SbcC"
                /gene="BMETHMGA300434c"
CDS             complement(509671..510759)
                /note="exonuclease SbcCD, D subunit"
                /gene="BMETHMGA300435c"
```

```
CDS             complement(511532..512416)
                /note="hypothetical protein"
                /gene="BMETHMGA300436c"
CDS             complement(512653..513456)
                /note="hypothetical protein"
                /gene="BMETHMGA300437c"
CDS             513906..514331
                /note="hypothetical protein"
                /gene="BMETHMGA300438"
CDS             complement(514501..514989)
                /note="hypothetical protein"
                /gene="BMETHMGA300439c"
CDS             complement(515172..515624)
                /note="two-component response regulator"
                /gene="BMETHMGA300440c"
CDS             complement(515786..516514)
                /note="cytochrome c biogenesis protein tansmembrane
                protein"
                /gene="BMETHMGA300441c"
CDS             complement(517086..518894)
                /note="ABC transporter"
                /gene="BMETHMGA300442c"
CDS             complement(518884..520656)
                /note="ABC transporter"
                /gene="BMETHMGA300443c"
CDS             521401..522348
                /note="transcriptional regulator Spx"
                /gene="BMETHMGA300444"
CDS             complement(522889..523263)
                /note="hypothetical protein"
                /gene="BMETHMGA300445c"
CDS             complement(523541..525457)
                /note="transketolase"
                /gene="BMETHMGA300446c"
CDS             complement(526184..526852)
                /note="resolvase domain protein"
                /gene="BMETHMGA300447c"
CDS             complement(526856..527179)
                /note="cell division supressor protein"
                /gene="BMETHMGA300448c"
CDS             527336..527959
                /note="LexA repressor"
                /gene="BMETHMGA300449"
CDS             528261..528587
                /note="hypothetical protein"
                /gene="BMETHMGA300450"
CDS             528817..529524
                /note="conserved membrane protein"
                /gene="BMETHMGA300451"
CDS             complement(529483..529794)
                /note="putative methyltransferase"
                /gene="BMETHMGA300452c"
CDS             complement(531197..531748)
```

```
                /note="hypothetical protein"
                /gene="BMETHMGA300454c"
CDS             532033..532359
                /note="hypothetical protein"
                /gene="BMETHMGA300455"
CDS             532455..533153
                /note="carbonic anhyddrase"
                /gene="BMETHMGA300456"
CDS             534025..534300
                /note="azoreductase"
                /gene="BMETHMGA300458"
CDS             complement(534740..535801)
                /note="sulfite oxidase"
                /gene="BMETHMGA300459c"
CDS             536220..536507
                /note="glucose-methanol-choline oxidoreductase"
                /gene="BMETHMGA300460"
CDS             536598..537104
                /note="glucose-methanol-choline oxidoreductase"
                /gene="BMETHMGA300461"
CDS             complement(538227..538703)
                /note="transposase"
                /gene="BMETHMGA300463c"
CDS             539584..540156
                /note="Appr-1-p processing protein"
                /gene="BMETHMGA300464"
CDS             complement(540501..541085)
                /note="lysince decarboxylase"
                /gene="BMETHMGA300465c"
CDS             complement(542290..543564)
                /note="cytochrome P450"
                /gene="BMETHMGA300466c"
fasta_record    544125..643528
                /note="ConsensusfromContig18 Average coverage:
       557.01"
                /label=ConsensusfromContig18 Average coverage: 557.01
                /colour=11
CDS             complement(544212..545423)
                /note="ABC transporter"
                /gene="BMETHMGA300467c"
CDS             complement(545416..546156)
                /note="ABC transporter"
                /gene="BMETHMGA300468c"
CDS             546765..547193
                /note="cell-cycle regulation protein- Hit"
                /gene="BMETHMGA300469"
CDS             complement(547802..548068)
                /note="transposase"
                /gene="BMETHMGA300470c"
CDS             548313..549431
                /note="phosphoserine aminotransferase"
                /gene="BMETHMGA300471"
CDS             549544..550080
```

```
                     /note="tryptophan transporter"
                     /gene="BMETHMGA300472"
     CDS             550590..550949
                     /note="hypothetical protein"
                     /gene="BMETHMGA300473"
     CDS             551206..551781
                     /note="transcriptional regulator MarR family"
                     /gene="BMETHMGA300474"
     CDS             552719..553579
                     /note="post-translocation molecular chaperone"
                     /gene="BMETHMGA300475"
     CDS             complement(554105..555049)
                     /note="3'-5' exoribonuclease"
                     /gene="BMETHMGA300476c"
     CDS             complement(555102..558089)
                     /note="hypothetical protein"
                     /gene="BMETHMGA300477c"
     CDS             complement(558086..559309)
                     /note="DNA repair exonuclease"
                     /gene="BMETHMGA300478c"
     CDS             complement(559454..560707)
                     /note="Na+ABC efflux transporter"
                     /gene="BMETHMGA300479c"
     CDS             complement(560700..561617)
                     /note="ABC transporter"
                     /gene="BMETHMGA300480c"
     CDS             complement(562167..562988)
                     /note="enoyl-CoA hydratase"
                     /gene="BMETHMGA300481c"
     CDS             complement(563160..564782)
                     /note="medium-chain fatty acid-CoA ligase"
                     /gene="BMETHMGA300482c"
     CDS             complement(564932..566350)
                     /note="coproporphyrinogen III oxidase"
                     /gene="BMETHMGA300483c"
     CDS             complement(566611..567480)
                     /note="HAD-superfamily hydrolase"
                     /gene="BMETHMGA300484c"
     CDS             complement(567624..567947)
                     /note="hypothetical protein"
                     /gene="BMETHMGA300485c"
     CDS             568175..569554
                     /note="glutathione synthetase ATP-binding protein"
                     /gene="BMETHMGA300486"
     CDS             569554..570654
                     /note="hypothetical protein"
                     /gene="BMETHMGA300487"
     CDS             570584..571993
                     /note="glutathione synthetase ATP-binding domain"
                     /gene="BMETHMGA300488"
     CDS             571990..573168
                     /note="hypothetical protein"
                     /gene="BMETHMGA300489"
```

```
CDS             573542..574465
                /note="putative transcriptional regulator"
                /gene="BMETHMGA300490"
CDS             574471..575742
                /note="multiple sugar binding transporter ATP-binding
                protein"
                /gene="BMETHMGA300491"
CDS             complement(576585..577964)
                /note="fumarate hydratase"
                /gene="BMETHMGA300492c"
CDS             complement(578366..579586)
                /note="isoaspartyl dipeptidase"
                /gene="BMETHMGA300493c"
CDS             complement(579653..581710)
                /note="penicillin-binding protein"
                /gene="BMETHMGA300494c"
CDS             complement(581911..582330)
                /note="hypothetical protein"
                /gene="BMETHMGA300495c"
CDS             complement(582669..584675)
                /note="Fe2+ transport system protein B"
                /gene="BMETHMGA300496c"
CDS             585187..586083
                /note="cation efflux family protein"
                /gene="BMETHMGA300497"
CDS             586265..587407
                /note="sensor histidine kinase"
                /gene="BMETHMGA300498"
CDS             587583..588152
                /note="ribosomal-protein serine acetyltransferase"
                /gene="BMETHMGA300499"
CDS             588306..589610
                /note="putative transporter or sensor"
                /gene="BMETHMGA300500"
CDS             589860..590243
                /note="transcriptional regulator, MerR"
                /gene="BMETHMGA300501"
CDS             590263..591654
                /note="2-oxo acid dehydrogenase"
                /gene="BMETHMGA300502"
CDS             complement(591803..593755)
                /note="sulfatase/glycerol phosphate lipoteichoic acid
                synthase 2"
                /gene="BMETHMGA300503c"
CDS             complement(594035..595444)
                /note="histidine kinase"
                /gene="BMETHMGA300504c"
CDS             complement(595434..596141)
                /note="two-component response regulator"
                /gene="BMETHMGA300505c"
CDS             complement(596580..597839)
                /note="serine protease HtrA"
                /gene="BMETHMGA300506c"
```

```
CDS             598351..599157
                /note="permease"
                /gene="BMETHMGA300507"
CDS             599150..600004
                /note="hypothetical protein"
                /gene="BMETHMGA300508"
CDS             complement(600321..601094)
                /note="Cof-like hydrolase"
                /gene="BMETHMGA300509c"
CDS             complement(601367..602269)
                /note="fructokinase"
                /gene="BMETHMGA300510c"
CDS             complement(605415..607154)
                /note="thiol reductant ABC exporter, CydC subunit"
                /gene="BMETHMGA300512c"
CDS             complement(607154..608869)
                /note="ABC transporter, ATP-binding protein"
                /gene="BMETHMGA300513c"
CDS             complement(608869..609882)
                /note="cytochrome d ubiquinol oxidase"
                /gene="BMETHMGA300514c"
CDS             complement(609884..611317)
                /note="cytochrome d ubiquinol oxidase, subunit I"
                /gene="BMETHMGA300515c"
CDS             611953..614046
                /note="penicillin-binding protein A"
                /gene="BMETHMGA300516"
CDS             complement(614148..615659)
                /note="SpoVR"
                /gene="BMETHMGA300517c"
CDS             complement(615688..616050)
                /note="hypothetical protein"
                /gene="BMETHMGA300518c"
CDS             complement(616538..618316)
                /note="glucocse-1,6-biphosphate synthase"
                /gene="BMETHMGA300520c"
CDS             618479..618913
                /note="disulfate oxidoreductase"
                /gene="BMETHMGA300521"
CDS             618915..619385
                /note="thioredoxin-like protein"
                /gene="BMETHMGA300522"
CDS             complement(619428..619889)
                /note="hypothetical protein"
                /gene="BMETHMGA300523c"
CDS             620103..621026
                /note="pseudouridine synthase"
                /gene="BMETHMGA300524"
CDS             complement(621179..621439)
                /note="transglycosylase protein"
                /gene="BMETHMGA300525c"
CDS             621735..622307
                /note="FMN reductase, NADPH-dependent"
```

```
                     /gene="BMETHMGA300526"
CDS                  622436..623053
                     /note="putative phosphate transport regulator"
                     /gene="BMETHMGA300527"
CDS                  623066..624064
                     /note="PiT family inorganic phosphate transporter"
                     /gene="BMETHMGA300528"
CDS                  624398..627427
                     /note="AcrB/AcrD/AcrF family transporter"
                     /gene="BMETHMGA300529"
CDS                  complement(627529..628560)
                     /note="sporulation integral membrane protein"
                     /gene="BMETHMGA300530c"
CDS                  628681..629319
                     /note="3-methyladenine DNA glycosylase"
                     /gene="BMETHMGA300531"
CDS                  complement(629341..630510)
                     /note="sporulation protein"
                     /gene="BMETHMGA300532c"
CDS                  complement(630855..632750)
                     /note="serine protein kinase"
                     /gene="BMETHMGA300533c"
CDS                  complement(633116..633865)
                     /note="nitroreductase"
                     /gene="BMETHMGA300534c"
CDS                  complement(633919..634263)
                     /note="RNA methyltransferase"
                     /gene="BMETHMGA300535c"
CDS                  complement(634507..635412)
                     /note="hypothetical protein"
                     /gene="BMETHMGA300536c"
CDS                  complement(635481..636668)
                     /note="iron-sulfur cluster-binding protein"
                     /gene="BMETHMGA300537c"
CDS                  636819..637388
                     /note="hypothetical protein"
                     /gene="BMETHMGA300538"
CDS                  complement(637953..638597)
                     /note="phosphate ABC transporter"
                     /gene="BMETHMGA300539c"
CDS                  638671..639096
                     /note="phosphate ABC transporter"
                     /gene="BMETHMGA300540"
CDS                  complement(639374..640240)
                     /note="phosphate binding protein"
                     /gene="BMETHMGA300541c"
CDS                  complement(640266..641972)
                     /note="hypothetical protein"
                     /gene="BMETHMGA300542c"
CDS                  641997..642446
                     /note="hypothetical protein"
                     /gene="BMETHMGA300543"
CDS                  complement(642748..643173)
```

```
                /note="hypothetical protein"
                /gene="BMETHMGA300544c"
fasta_record    643529..729456
                /note="ConsensusfromContig19 Average coverage:
       575.27"
                /label=ConsensusfromContig19 Average coverage: 575.27
                /colour=10
CDS             complement(643603..645660)
                /note="putative cell division protein FtsA"
                /gene="BMETHMGA300545c"
CDS             646009..647085
                /note="3-deoxy-7-phosphoheptulonate synthase"
                /gene="BMETHMGA300546"
CDS             647449..648447
                /note="catabolite control protein A"
                /gene="BMETHMGA300547"
CDS             complement(648563..649750)
                /note="acetoin utilization protein "
                /gene="BMETHMGA300548c"
CDS             complement(649720..650367)
                /note="acetoin dehydrogenase"
                /gene="BMETHMGA300549c"
CDS             complement(650545..651177)
                /note="acetoin utilization"
                /gene="BMETHMGA300550c"
CDS             651981..653699
                /note="acetyl-CoA synthetase"
                /gene="BMETHMGA300551"
CDS             complement(653752..656715)
                /note="peptidoglycan glycosyltransferase"
                /gene="BMETHMGA300552c"
CDS             657060..658328
                /note="tyrosyl-tRNA synthetase"
                /gene="BMETHMGA300553"
CDS             complement(658406..659557)
                /note="aldehyde dehydrogenase"
                /gene="BMETHMGA300554c"
CDS             complement(660181..660852)
                /note="30S ribosomal protein S4"
                /gene="BMETHMGA300555c"
CDS             661286..663178
                /note="diguanylate cyclase"
                /gene="BMETHMGA300556"
CDS             complement(663275..663790)
                /note="hypothetical protein"
                /gene="BMETHMGA300557c"
CDS             663886..664542
                /note="transcriptional regulator"
                /gene="BMETHMGA300558"
CDS             complement(664565..665188)
                /note="histidinol-phosphatase"
                /gene="BMETHMGA300559c"
CDS             665539..667227
```

```
                     /note="septation ring formation regulator EzrA"
                     /gene="BMETHMGA300560"
     CDS             667429..668571
                     /note="cysteine desulfurase"
                     /gene="BMETHMGA300561"
     CDS             668573..669781
                     /note="thiamine biosynthesis/tRNA modification
            protein"
                     /gene="BMETHMGA300562"
     CDS             670335..671921
                     /note="acetyl-CoA synthetase/acetate-CoA ligase"
                     /gene="BMETHMGA300563"
     CDS             672192..673124
                     /note="chloroamphenicol-sensitive protein RarD"
                     /gene="BMETHMGA300564"
     CDS             complement(673315..674115)
                     /note="inorganic polyphosphate/ATP-NAD kinase"
                     /gene="BMETHMGA300565c"
     CDS             674267..675277
                     /note="signal peptide peptidase"
                     /gene="BMETHMGA300566"
     CDS             675292..675864
                     /note="RDD domain containing protein"
                     /gene="BMETHMGA300567"
     CDS             675942..676646
                     /note="hypothetical protein"
                     /gene="BMETHMGA300568"
     CDS             676664..677149
                     /note="sporulatio protein YtfJ"
                     /gene="BMETHMGA300569"
     CDS             677279..677779
                     /note="thioredoxin peroxidase"
                     /gene="BMETHMGA300570"
     CDS             678153..678881
                     /note="adenine-specific methyltransferase"
                     /gene="BMETHMGA300571"
     CDS             679169..680359
                     /note="acetate/propionate kinase"
                     /gene="BMETHMGA300572"
     CDS             680474..681316
                     /note="EcsC protein"
                     /gene="BMETHMGA300573"
     CDS             681313..681825
                     /note="molybdenum cofactor synthesis protein"
                     /gene="BMETHMGA300574"
     CDS             complement(681869..682183)
                     /note="hypothetical protein"
                     /gene="BMETHMGA300575c"
     CDS             682712..683698
                     /note="argininosuccinate synthase"
                     /gene="BMETHMGA300576"
     CDS             683824..685089
                     /note="argininosuccinate lyase"
```

```
                     /gene="BMETHMGA300577"
     CDS             685355..685921
                     /note="CBS domain protein"
                     /gene="BMETHMGA300578"
     CDS             686317..687531
                     /note="alanine dehydrogenase"
                     /gene="BMETHMGA300579"
     CDS             complement(687631..688728)
                     /note="xaa-Pro peptidase"
                     /gene="BMETHMGA300580c"
     CDS             688895..689575
                     /note="metal-dependent hydrolase"
                     /gene="BMETHMGA300581"
     CDS             690375..691385
                     /note="GntR family transcriptional regulator"
                     /gene="BMETHMGA300582"
     CDS             complement(691544..691849)
                     /note="hypothetical protein"
                     /gene="BMETHMGA300583c"
     CDS             692048..692986
                     /note="3'-5'-biphosphate nucleotidase"
                     /gene="BMETHMGA300584"
     CDS             complement(693148..693651)
                     /note="sporulation protein"
                     /gene="BMETHMGA300585c"
     CDS             complement(693648..693998)
                     /note="hypothetical protein"
                     /gene="BMETHMGA300586c"
     CDS             694139..697552
                     /note="DNA polymerase III, alpha subunit"
                     /gene="BMETHMGA300587"
     CDS             697710..698921
                     /note="malate dehydrogenase"
                     /gene="BMETHMGA300588"
     CDS             698991..699554
                     /note="GntR family transcriptional regulator"
                     /gene="BMETHMGA300589"
     CDS             699840..700625
                     /note="acetyl-CoA carboxylase beta subunit"
                     /gene="BMETHMGA300590"
     CDS             700610..701587
                     /note="acetyl-CoA carboxylase, alpha subunit"
                     /gene="BMETHMGA300591"
     CDS             701797..702756
                     /note="6-phosphofructokinase"
                     /gene="BMETHMGA300592"
     CDS             702794..704554
                     /note="pyruvate kinase"
                     /gene="BMETHMGA300593"
     CDS             704569..705057
                     /note="FxsA cytoplasmic membrane protein"
                     /gene="BMETHMGA300594"
     CDS             complement(705173..706120)
```

```
                /note="spurulation integral membrane protein"
                /gene="BMETHMGA300595c"
CDS             706559..706951
                /note="Hypothetical protein"
                /gene="BMETHMGA300596"
CDS             707509..708627
                /note="citrate synthase"
                /gene="BMETHMGA300597"
CDS             708887..710155
                /note="isocitrate dehydrogenase"
                /gene="BMETHMGA300598"
CDS             710289..711227
                /note="malate dehydrogenase"
                /gene="BMETHMGA300599"
CDS             711448..712164
                /note="PhoP, two-component response regulator"
                /gene="BMETHMGA300600"
CDS             712161..713927
                /note="two-component sensor histidine kinase"
                /gene="BMETHMGA300601"
CDS             714177..715142
                /note="protease specific for phage lambda cII
       repressor"
                /gene="BMETHMGA300602"
CDS             715135..716067
                /note="protease specific for phage lambda cII
       repressor"
                /gene="BMETHMGA300603"
CDS             716304..718820
                /note="DNA polymerase I"
                /gene="BMETHMGA300604"
CDS             718945..719673
                /note="formamidopyrimidine-DNA glycosylase"
                /gene="BMETHMGA300605"
CDS             719750..720385
                /note="sporulation protein YtaF"
                /gene="BMETHMGA300606"
CDS             720406..721008
                /note="dephospho-CoA kinase"
                /gene="BMETHMGA300607"
CDS             721244..722275
                /note="glyceraldehyde-3-phosphate dehydroganse"
                /gene="BMETHMGA300608"
CDS             722596..722976
                /note="S-adenosylmethionine decarboxylase proenzyme"
                /gene="BMETHMGA300609"
CDS             723185..723655
                /note="NrdR family transcriptional regulator"
                /gene="BMETHMGA300610"
CDS             723716..725110
                /note="replication initiation and membrane
       attachement
                protein"
```

```
                 /gene="BMETHMGA300611"
CDS              725127..726056
                 /note="primosomal protein DnaI"
                 /gene="BMETHMGA300612"
CDS              726203..727069
                 /note="sporulation protein YtxC"
                 /gene="BMETHMGA300613"
CDS              727438..729372
                 /note="threonyl-tRNA synthetase"
                 /gene="BMETHMGA300614"
fasta_record     729457..753176
                 /note="ConsensusfromContig20 Average coverage:
       607.14"
                 /label=ConsensusfromContig20 Average coverage: 607.14
                 /colour=11
CDS              complement(730204..730704)
                 /note="acetone carboxylase, gamma subunit"
                 /gene="BMETHMGA300615c"
CDS              complement(730737..733004)
                 /note="acetone carboxylase, alpha subunit"
                 /gene="BMETHMGA300616c"
CDS              complement(733028..735175)
                 /note="acetone carboyxlase, beta subunit"
                 /gene="BMETHMGA300617c"
CDS              735555..737525
                 /note="transcriptional activator of acetoin glycerol
                 metabolism"
                 /gene="BMETHMGA300618"
CDS              complement(737800..738255)
                 /note="OsmC family protein"
                 /gene="BMETHMGA300619c"
CDS              complement(738503..739630)
                 /note="lysine 2,3-aminomutase"
                 /gene="BMETHMGA300620c"
CDS              complement(739909..740331)
                 /note="high-affinity nickel transporter"
                 /gene="BMETHMGA300621c"
CDS              complement(741067..742110)
                 /note="cytathionine gamma-synthase"
                 /gene="BMETHMGA300622c"
CDS              743047..744315
                 /note="aldehyde dehyrogenase"
                 /gene="BMETHMGA300623"
CDS              complement(745478..745900)
                 /note="GtrA family protein"
                 /gene="BMETHMGA300624c"
CDS              complement(745900..746901)
                 /note="glycosyl transferase family 2"
                 /gene="BMETHMGA300625c"
CDS              complement(746933..748618)
                 /note="hypothetical protein"
                 /gene="BMETHMGA300626c"
CDS              complement(749044..750594)
```

```
                     /note="acyl-CoA synthetase"
                     /gene="BMETHMGA300627c"
     CDS             751001..751525
                     /note="hypothetical protein"
                     /gene="BMETHMGA300628"
     CDS             751804..752448
                     /note="ABC transporter"
                     /gene="BMETHMGA300629"
     fasta_record    753177..960522
                     /note="ConsensusfromContig21 Average coverage:
            531.26"
                     /label=ConsensusfromContig21 Average coverage: 531.26
                     /colour=10
     CDS             753515..755317
                     /note="DNA primase"
                     /gene="BMETHMGA300630"
     CDS             755360..756487
                     /note="RNA polymerase sigma factor RpoD"
                     /gene="BMETHMGA300631"
     CDS             757083..758234
                     /note="acyl-CoA dehydrogenase"
                     /gene="BMETHMGA300632"
     CDS             758422..758793
                     /note="cytochrome c-550"
                     /gene="BMETHMGA300633"
     CDS             758961..759674
                     /note="hypothetical protein"
                     /gene="BMETHMGA300634"
     CDS             759671..760786
                     /note="hypothetical protein"
                     /gene="BMETHMGA300635"
     CDS             complement(760828..761781)
                     /note="4-hydroxy-3-methylbut-2-enyl disphosphate
                     reductase"
                     /gene="BMETHMGA300636c"
     CDS             complement(761918..762508)
                     /note="hypothetical protein"
                     /gene="BMETHMGA300637c"
     CDS             762729..764039
                     /note="DEAD-box ATP dependent DNA helicase"
                     /gene="BMETHMGA300638"
     CDS             764036..764947
                     /note="apurinic endonuclease IV"
                     /gene="BMETHMGA300639"
     CDS             765472..766353
                     /note="hypothetical protein"
                     /gene="BMETHMGA300640"
     CDS             766453..767271
                     /note="Zinc uptake system ATP-binding protein"
                     /gene="BMETHMGA300641"
     CDS             767268..768119
                     /note="high affinity zinc uptake system membrane
            protein"
```

```
                     /gene="BMETHMGA300642"
CDS                  768125..768535
                     /note="zinc-specific metalloregulatory protein"
                     /gene="BMETHMGA300643"
CDS                  complement(768628..769209)
                     /note="hypothetical protein"
                     /gene="BMETHMGA300644c"
CDS                  769401..769766
                     /note="hypothetical protein"
                     /gene="BMETHMGA300645"
CDS                  complement(769826..770848)
                     /note="4-hydroxy-3-methylbut-2-en-l-yl-diphosphate
                     syntase"
                     /gene="BMETHMGA300646c"
CDS                  771366..772694
                     /note="major facilitator family transporter"
                     /gene="BMETHMGA300647"
CDS                  complement(772797..773564)
                     /note="hypothetical protein"
                     /gene="BMETHMGA300648c"
CDS                  773770..774252
                     /note="hypothetical protein"
                     /gene="BMETHMGA300649"
CDS                  774535..775146
                     /note="superoxide dismutase"
                     /gene="BMETHMGA300650"
CDS                  775331..776872
                     /note="exopolyphosphatase"
                     /gene="BMETHMGA300651"
CDS                  776921..779044
                     /note="polyphosphate kinase"
                     /gene="BMETHMGA300652"
CDS                  779152..779505
                     /note="hypothetical protein"
                     /gene="BMETHMGA300653"
CDS                  780249..784565
                     /note="Cation-transporting ATPase"
                     /gene="BMETHMGA300655"
CDS                  785321..786547
                     /note="major facilitator superfamily transporter"
                     /gene="BMETHMGA300656"
CDS                  786790..788907
                     /note="penicillin-binding protein transpeptidase"
                     /gene="BMETHMGA300657"
CDS                  789542..790369
                     /note="phosphate ABC transporter ATPase"
                     /gene="BMETHMGA300658"
CDS                  790448..791107
                     /note="phosphate transport system regulatory protein"
                     /gene="BMETHMGA300659"
CDS                  complement(791138..791683)
                     /note="aminodeoxychorismate lyase"
                     /gene="BMETHMGA300660c"
```

```
CDS             complement(791646..792035)
                /note="hypothetical protein"
                /gene="BMETHMGA300661c"
CDS             792360..793040
                /note="hypothetical protein"
                /gene="BMETHMGA300662"
CDS             793044..795455
                /note="glycosyl transferase, group I"
                /gene="BMETHMGA300663"
CDS             795545..796138
                /note="hypothetical protein"
                /gene="BMETHMGA300664"
CDS             796412..796999
                /note="5-formyltetrahydrofolate cyclo-ligase"
                /gene="BMETHMGA300665"
CDS             797585..798997
                /note="Rhomboid family protein"
                /gene="BMETHMGA300666"
CDS             complement(798963..799838)
                /note="hypothetical protein"
                /gene="BMETHMGA300667c"
CDS             800284..801261
                /note="Glucokinase"
                /gene="BMETHMGA300668"
CDS             801550..803445
                /note="sulfatase"
                /gene="BMETHMGA300669"
CDS             803526..804785
                /note="gamma-D-glutamyl-L-diamino acid endopeptidase"
                /gene="BMETHMGA300670"
CDS             805065..806237
                /note="hypothetical protein"
                /gene="BMETHMGA300671"
CDS             complement(806532..807026)
                /note="hypothetical protein"
                /gene="BMETHMGA300672c"
CDS             807105..807740
                /note="Zn-dependent hydrolase"
                /gene="BMETHMGA300673"
CDS             807929..809038
                /note="hypothetical protein"
                /gene="BMETHMGA300674"
CDS             complement(809405..809935)
                /note="transcriptional regulator, ArsR family"
                /gene="BMETHMGA300675c"
CDS             810577..811548
                /note="competence protein ComGA"
                /gene="BMETHMGA300676"
CDS             811502..812554
                /note="comptence operon protein 2"
                /gene="BMETHMGA300677"
CDS             812603..812881
                /note="comG operon protein 3"
```

```
                     /gene="BMETHMGA300678"
     CDS             812878..813327
                     /note="ComG operon protein 4"
                     /gene="BMETHMGA300679"
     CDS             813302..813631
                     /note="ComG operon protein 5"
                     /gene="BMETHMGA300680"
     CDS             813645..814055
                     /note="ComG operon protein 6"
                     /gene="BMETHMGA300681"
     CDS             814022..814438
                     /note="ComG operon protein 7"
                     /gene="BMETHMGA300682"
     CDS             814651..814956
                     /note="shikimate kinase"
                     /gene="BMETHMGA300683"
     CDS             complement(815221..815889)
                     /note="hypothetical protein"
                     /gene="BMETHMGA300684c"
     CDS             complement(815981..817678)
                     /note="DNA/RNA helicase"
                     /gene="BMETHMGA300685c"
     CDS             818149..819270
                     /note="glycine cleavage sytem T protein"
                     /gene="BMETHMGA300686"
     CDS             819274..820620
                     /note="glycine dehydrdogenase subunit 1"
                     /gene="BMETHMGA300687"
     CDS             820613..822076
                     /note="glycine dehydrogenase subunit 2"
                     /gene="BMETHMGA300688"
     CDS             complement(822329..823135)
                     /note="hypothetical protein"
                     /gene="BMETHMGA300689c"
     CDS             complement(823483..823863)
                     /note="Rhodanese domain protein"
                     /gene="BMETHMGA300690c"
     CDS             824051..824887
                     /note="lipoate protein ligase"
                     /gene="BMETHMGA300691"
     CDS             825416..826450
                     /note="RNA-directed DNA polymerase"
                     /gene="BMETHMGA300692"
     CDS             826792..829353
                     /note="ribonucleoside-diphosphate reductase, alpha
                     subunit"
                     /gene="BMETHMGA300693"
     CDS             829585..830469
                     /note="phospholipase"
                     /gene="BMETHMGA300694"
     CDS             complement(830632..831024)
                     /note="Hypothetical protein YqhP"
                     /gene="BMETHMGA300695c"
```

```
CDS             complement(831082..832026)
                /note="hypothetical protein YqhQ"
                /gene="BMETHMGA300696c"
CDS             832153..832734
                /note="hypothetical protein YqhR"
                /gene="BMETHMGA300697"
CDS             832817..833278
                /note="3-dehydroquinate dehydratase"
                /gene="BMETHMGA300698"
CDS             833214..834326
                /note="xaa-Pro dipeptidase"
                /gene="BMETHMGA300699"
CDS             834353..834910
                /note="elongation factor P"
                /gene="BMETHMGA300700"
CDS             complement(835274..835942)
                /note="hypothetical protein"
                /gene="BMETHMGA300701c"
CDS             836286..837086
                /note="stage III sporulation protein AA"
                /gene="BMETHMGA300702"
CDS             837087..837602
                /note="stage III sporulation protein AB"
                /gene="BMETHMGA300703"
CDS             838268..839467
                /note="stage III sporulation protein AE"
                /gene="BMETHMGA300704"
CDS             839476..840108
                /note="stage III sporulation protein AF"
                /gene="BMETHMGA300705"
CDS             840098..840754
                /note="stage III sporulation protein AG"
                /gene="BMETHMGA300706"
CDS             840768..841337
                /note="stage III sporulation protein AH"
                /gene="BMETHMGA300707"
CDS             841474..841983
                /note="acetyl-CoA carboxylase, biotin carboxyl
       carrier
                protein"
                /gene="BMETHMGA300708"
CDS             841997..843355
                /note="acetyl-CoA carboxylase, biotin carboxylase"
                /gene="BMETHMGA300709"
CDS             843395..843778
                /note="hypothetical protein"
                /gene="BMETHMGA300710"
CDS             843966..844355
                /note="transcriptional antitermination factor NusB"
                /gene="BMETHMGA300711"
CDS             844397..845266
                /note="methylenetetrahydrofolate
                dehydrogenase/cyclohydrolase"
```

```
                     /gene="BMETHMGA300712"
     CDS             845397..846770
                     /note="exodeoxyribonuclease VII, large subunit"
                     /gene="BMETHMGA300713"
     CDS             847020..847868
                     /note="geranyltranstransferase"
                     /gene="BMETHMGA300714"
     CDS             848067..849962
                     /note="1-deoxy-D-xylulose-5-phosphate synthase"
                     /gene="BMETHMGA300715"
     CDS             849966..850808
                     /note="putative cell adhesion protein"
                     /gene="BMETHMGA300716"
     CDS             851031..851480
                     /note="arginine repressor"
                     /gene="BMETHMGA300717"
     CDS             851770..853188
                     /note="DNA repair protein RecN"
                     /gene="BMETHMGA300718"
     CDS             853565..854602
                     /note="stage IV sporulation protein B"
                     /gene="BMETHMGA300719"
     CDS             854974..855681
                     /note="stage 0 sporulation protein A"
                     /gene="BMETHMGA300720"
     CDS             855840..856565
                     /note="glycerophosphoryl diester phosphodiesterase"
                     /gene="BMETHMGA300721"
     CDS             complement(856624..856893)
                     /note="hypothetical protein"
                     /gene="BMETHMGA300722c"
     CDS             857053..859125
                     /note="sigma-54-dependent transcriptional activator"
                     /gene="BMETHMGA300723"
     CDS             859260..860162
                     /note="phosphate butyryltransferase"
                     /gene="BMETHMGA300724"
     CDS             860194..861288
                     /note="Butyrul-CoA dehydrogenase"
                     /gene="BMETHMGA300725"
     CDS             861504..862628
                     /note="butyrate kinase"
                     /gene="BMETHMGA300726"
     CDS             862856..864064
                     /note="dihydrolipoyl dehydrogenase"
                     /gene="BMETHMGA300727"
     CDS             864146..865141
                     /note="3-methyl-2-oxobutanoate dehydrogenase"
                     /gene="BMETHMGA300728"
     CDS             865154..866137
                     /note="n3-methyl-2-oxobutanoate dehydrogenase"
                     /gene="BMETHMGA300729"
     CDS             866199..867485
```

```
                     /note="dihydrolipoamide acetyltransferase"
                     /gene="BMETHMGA300730"
     CDS             867882..869762
                     /note="methylmalonyl-CoA mutase, small subunit"
                     /gene="BMETHMGA300731"
     CDS             869752..871947
                     /note="methylmalonyl-CoA mutase, large subunit"
                     /gene="BMETHMGA300732"
     CDS             871947..873074
                     /note="arginine/ornithine transport system ATPase"
                     /gene="BMETHMGA300733"
     CDS             873281..873754
                     /note="hypothetical protein"
                     /gene="BMETHMGA300734"
     CDS             874091..874528
                     /note="hypothetical protein"
                     /gene="BMETHMGA300735"
     CDS             874808..875602
                     /note="arginine-binding extracellular protein"
                     /gene="BMETHMGA300736"
     CDS             875697..876356
                     /note="arginine ABC transport permease"
                     /gene="BMETHMGA300737"
     CDS             876487..877071
                     /note="arginine transport ATP-binding protein"
                     /gene="BMETHMGA300738"
     CDS             877535..877957
                     /note="methylmalonyl-CoA epimerase"
                     /gene="BMETHMGA300739"
     CDS             877950..879500
                     /note="propionyl-CoA carboxylase, beta subunit"
                     /gene="BMETHMGA300740"
     CDS             879566..880684
                     /note="tripeptidase T-like protein"
                     /gene="BMETHMGA300741"
     CDS             880825..881310
                     /note="chemotaxis protein"
                     /gene="BMETHMGA300742"
     CDS             881397..882698
                     /note="DNA polymerase IV"
                     /gene="BMETHMGA300743"
     CDS             882821..884230
                     /note="6-phosphogluconate dehydrogenase"
                     /gene="BMETHMGA300744"
     CDS             884354..885274
                     /note="hypothetical protein"
                     /gene="BMETHMGA300745"
     CDS             complement(885348..886805)
                     /note="glucose-6-phosphate 1-dehydrogenase"
                     /gene="BMETHMGA300746c"
     CDS             887099..888031
                     /note="ribonuclease Z"
                     /gene="BMETHMGA300747"
```

```
CDS             complement(888106..889140)
                /note="NADPH dehydrogenase"
                /gene="BMETHMGA300748c"
CDS             complement(889304..890143)
                /note="pyrroline-5-carboxylate reductase"
                /gene="BMETHMGA300749c"
CDS             890677..891639
                /note="Zn-dependent hydrolase"
                /gene="BMETHMGA300750"
CDS             891654..892445
                /note="oxidoreductase"
                /gene="BMETHMGA300751"
CDS             893080..894339
                /note="UV-damage repair protein UvrX"
                /gene="BMETHMGA300752"
CDS             894363..894710
                /note="hypothetical protein"
                /gene="BMETHMGA300753"
CDS             894822..895151
                /note="hypothetical protein"
                /gene="BMETHMGA300754"
CDS             895445..896776
                /note="cytochrome bd-type ubiquinol oxidase subunit"
                /gene="BMETHMGA300755"
CDS             896799..897815
                /note="cytochrome d ubiquinol oxidase, subunit II"
                /gene="BMETHMGA300756"
CDS             complement(898005..898976)
                /note="alpha/beta hydrolase"
                /gene="BMETHMGA300757c"
CDS             complement(899274..900200)
                /note="hydroxymethylglutaryl-CoA lyase"
                /gene="BMETHMGA300758c"
CDS             900508..901197
                /note="3-oxoacid CoA-transferase, subunit A"
                /gene="BMETHMGA300759"
CDS             901220..901888
                /note="3-oxoacid CoA-transferase, subunit B"
                /gene="BMETHMGA300760"
CDS             complement(902076..902516)
                /note="nitric oxide sensitive transcriptional
       regulator"
                /gene="BMETHMGA300761c"
CDS             902761..903957
                /note="nitric oxide dioxygenase"
                /gene="BMETHMGA300762"
CDS             complement(904587..905477)
                /note="aldo/keto reductase"
                /gene="BMETHMGA300763c"
CDS             905746..906303
                /note="methanol dehydrogenase activator protein"
                /gene="BMETHMGA300764"
CDS             906300..907463
```

```
                     /note="hypothetical protein"
                     /gene="BMETHMGA300765"
     CDS             907626..908267
                     /note="Stage II sporulation protein M"
                     /gene="BMETHMGA300766"
     CDS             908412..908894
                     /note="Fe2+/Zn2+ uptake regulation protein"
                     /gene="BMETHMGA300767"
     CDS             909239..910150
                     /note="tyrosine recombinase xerD"
                     /gene="BMETHMGA300768"
     CDS             910364..911548
                     /note="phosphopentomutase"
                     /gene="BMETHMGA300769"
     CDS             911548..912384
                     /note="purine nucleoside phosphorylase"
                     /gene="BMETHMGA300770"
     CDS             912411..913715
                     /note="pyrimidine-nucleoside phosphorylase"
                     /gene="BMETHMGA300771"
     CDS             913834..915018
                     /note="D-alanyl-D-alanine carboxypeptidase"
                     /gene="BMETHMGA300772"
     CDS             915208..915618
                     /note="anti-sigma F factor antagonist"
                     /gene="BMETHMGA300773"
     CDS             915621..916061
                     /note="anti-sigma F factor"
                     /gene="BMETHMGA300774"
     CDS             916074..916832
                     /note="sporulation sigma factor SigF"
                     /gene="BMETHMGA300775"
     CDS             916922..917605
                     /note="stage V sporulation protein AA"
                     /gene="BMETHMGA300776"
     CDS             917595..918017
                     /note="stage V sporulation protein AB"
                     /gene="BMETHMGA300777"
     CDS             918070..919554
                     /note="stage V sporulation protein AF"
                     /gene="BMETHMGA300778"
     CDS             919746..921077
                     /note="diaminopimelate decarboxylase"
                     /gene="BMETHMGA300779"
     CDS             921444..921806
                     /note="N-acetyltransferase GCN5"
                     /gene="BMETHMGA300780"
     CDS             complement(921810..922337)
                     /note="hypothetical protein"
                     /gene="BMETHMGA300781c"
     CDS             922746..923513
                     /note="chromosome segregation and condensation
          protein A"
```

```
                 /gene="BMETHMGA300782"
CDS              923671..924084
                 /note="chromosome segregation and condensation
       protein B"
                 /gene="BMETHMGA300783"
CDS              924288..924776
                 /note="hypothetical protein"
                 /gene="BMETHMGA300784"
CDS              924907..925797
                 /note="superoixde dismutase"
                 /gene="BMETHMGA300785"
CDS              926102..927250
                 /note="D-alanyl-D-alanine carboxypeptidase"
                 /gene="BMETHMGA300786"
CDS              927243..927836
                 /note="spore maturation protein"
                 /gene="BMETHMGA300787"
CDS              927833..928375
                 /note="spore maturation protein B"
                 /gene="BMETHMGA300788"
CDS              928520..929239
                 /note="pseudouridine synthase"
                 /gene="BMETHMGA300789"
CDS              929416..929940
                 /note="thiol-disulfide oxidoreductase"
                 /gene="BMETHMGA300790"
CDS              929961..931589
                 /note="cytochrome c biogenesis protein"
                 /gene="BMETHMGA300791"
CDS              931847..932782
                 /note="cytochrome c biogenesis protein"
                 /gene="BMETHMGA300792"
CDS              932884..933600
                 /note="response regulator protein"
                 /gene="BMETHMGA300793"
CDS              933597..935378
                 /note="two-component histidine kinase"
                 /gene="BMETHMGA300794"
CDS              935614..936165
                 /note="RNA polymerase sigma-70 factor"
                 /gene="BMETHMGA300795"
CDS              936281..937384
                 /note="hypothetical protein"
                 /gene="BMETHMGA300796"
CDS              937465..938127
                 /note="peptidase M23"
                 /gene="BMETHMGA300797"
CDS              938290..939102
                 /note="histidinol phosphate phosphatase"
                 /gene="BMETHMGA300798"
CDS              complement(939167..940747)
                 /note="D-3-phosphoglycerate dehydrogenase"
                 /gene="BMETHMGA300799c"
```

```
CDS             941301..941912
                /note="hypothetical protein"
                /gene="BMETHMGA300800"
CDS             942560..943624
                /note="hypothetical protein"
                /gene="BMETHMGA300801"
CDS             943615..945108
                /note="ATP-dependent DNA helicase"
                /gene="BMETHMGA300802"
CDS             945122..945733
                /note="hypothetical protein"
                /gene="BMETHMGA300803"
CDS             945726..946385
                /note="peptidoglycan-binding lysin domain protein"
                /gene="BMETHMGA300804"
CDS             946504..946962
                /note="hypothetical protein"
                /gene="BMETHMGA300805"
CDS             947176..947940
                /note="phosphoesterase"
                /gene="BMETHMGA300806"
CDS             948088..948681
                /note="adaptor protein mecA 2"
                /gene="BMETHMGA300807"
CDS             949058..950335
                /note="glutamate dehydrogenase"
                /gene="BMETHMGA300808"
CDS             950491..951468
                /note="FAD-dependent pyridine nucleotide disulfide
                oxidoreductase"
                /gene="BMETHMGA300809"
CDS             951465..952436
                /note="asparaginase/glutaminase"
                /gene="BMETHMGA300810"
CDS             952574..953263
                /note="hypothetical protein"
                /gene="BMETHMGA300811"
CDS             953359..954195
                /note="sporulation specific n-acetylmuramoyl-L-
       alanine
                amidase"
                /gene="BMETHMGA300812"
CDS             954394..955557
                /note="germination protein"
                /gene="BMETHMGA300813"
CDS             955742..956404
                /note="glycosyltransferase"
                /gene="BMETHMGA300814"
CDS             956778..957458
                /note="cytidylate kinase"
                /gene="BMETHMGA300815"
CDS             957455..958036
```

```
                   /note="1-acylglycerol-3-phosphate O-
        acetyltransferase"
                   /gene="BMETHMGA300816"
CDS                958189..959328
                   /note="RNA binding S1 domain protein"
                   /gene="BMETHMGA300817"
CDS                959398..960396
                   /note="isopentenyl pyrophosphate isomerase"
                   /gene="BMETHMGA300818"
fasta_record       960523..1060158
                   /note="ConsensusfromContig22 Average coverage:
        629.45"
                   /label=ConsensusfromContig22 Average coverage: 629.45
                   /colour=11
CDS                961031..963310
                   /note="hypothetical protein"
                   /gene="BMETHMGA300819"
CDS                complement(963594..964151)
                   /note="NADPH dependent FMN reductase"
                   /gene="BMETHMGA300820c"
CDS                complement(964159..965781)
                   /note="nitrilase/cyanide hydratase and apolipoprotein
                   N-acetyltransferase"
                   /gene="BMETHMGA300821c"
CDS                complement(965723..966469)
                   /note="thiamine-monophosphate kinase"
                   /gene="BMETHMGA300822c"
CDS                complement(966711..967877)
                   /note="glycosyl transferase, group 1"
                   /gene="BMETHMGA300823c"
CDS                complement(967923..968855)
                   /note="methylenetetrahydrofolate reductase"
                   /gene="BMETHMGA300824c"
CDS                complement(968887..969405)
                   /note="hypothetical protein"
                   /gene="BMETHMGA300825c"
CDS                complement(969456..970427)
                   /note="amidohydrolase 2"
                   /gene="BMETHMGA300826c"
CDS                complement(970366..970872)
                   /note="N-acetyltransferase"
                   /gene="BMETHMGA300827c"
CDS                complement(970891..972018)
                   /note="radical SAM protein"
                   /gene="BMETHMGA300828c"
CDS                972267..973706
                   /note="transcriptional regulator, GntR family"
                   /gene="BMETHMGA300829"
CDS                974834..975649
                   /note="taurine ABC transporter permease"
                   /gene="BMETHMGA300830"
CDS                975675..976730
                   /note="aliphatic sulfonate ABC transporter"
```

```
                     /gene="BMETHMGA300831"
     CDS             976753..977607
                     /note="taurine ABC transporter ATP-binding protein"
                     /gene="BMETHMGA300832"
     CDS             977655..978593
                     /note="taurine dioxygenase"
                     /gene="BMETHMGA300833"
     CDS             978693..979064
                     /note="hypothetical protein"
                     /gene="BMETHMGA300834"
     CDS             complement(979061..979732)
                     /note="short-chain dehydrogenase/reductase SDR"
                     /gene="BMETHMGA300835c"
     CDS             981692..982474
                     /note="hypothetical protein"
                     /gene="BMETHMGA300837"
     CDS             complement(982685..984058)
                     /note="mannose-6-phosphate isomerase"
                     /gene="BMETHMGA300838c"
     CDS             complement(984181..984633)
                     /note="urease accessory protein"
                     /gene="BMETHMGA300839c"
     CDS             complement(984835..985650)
                     /note="urease accessory protein UreD"
                     /gene="BMETHMGA300840c"
     CDS             complement(985647..986267)
                     /note="urease accessory protein UreG"
                     /gene="BMETHMGA300841c"
     CDS             complement(986288..986962)
                     /note="urease accessory protein UreF"
                     /gene="BMETHMGA300842c"
     CDS             complement(986952..987401)
                     /note="urease accessory protein UreE"
                     /gene="BMETHMGA300843c"
     CDS             complement(987414..989150)
                     /note="urease subunit alpha"
                     /gene="BMETHMGA300844c"
     CDS             complement(989128..989454)
                     /note="urease subunit B"
                     /gene="BMETHMGA300845c"
     CDS             complement(989470..989772)
                     /note="urease subunit gamma"
                     /gene="BMETHMGA300846c"
     CDS             complement(989790..990491)
                     /note="urea ABC transporter ATP-binding protein"
                     /gene="BMETHMGA300847c"
     CDS             complement(990485..991234)
                     /note="urea ABC transporter ATP-binding protein"
                     /gene="BMETHMGA300848c"
     CDS             complement(991209..992264)
                     /note="urea ABC transporter permease"
                     /gene="BMETHMGA300849c"
     CDS             complement(992278..993261)
```

```
                 /note="urea ABC transporter permease"
                 /gene="BMETHMGA300850c"
CDS              complement(993240..994496)
                 /note="urea ABC transporter, urea binding protein"
                 /gene="BMETHMGA300851c"
CDS              995792..996223
                 /note="transposase"
                 /gene="BMETHMGA300852"
CDS              complement(996839..1000885)
                 /note="S-layer domain protein"
                 /gene="BMETHMGA300853c"
CDS              complement(1001256..1002227)
                 /note="UDP-glucose 4-epimerase"
                 /gene="BMETHMGA300854c"
CDS              complement(1002422..1004701)
                 /note="2'3'-cyclic nucleotide
                 2'phosphodiesterase/3'nucletidase bifunctional
       periplasmic
                 protein"
                 /gene="BMETHMGA300855c"
CDS              complement(1005764..1006540)
                 /note="hypothetical protein"
                 /gene="BMETHMGA300856c"
CDS              complement(1007163..1008536)
                 /note="mannose-6-phoshphate isomerase"
                 /gene="BMETHMGA300857c"
CDS              complement(1009258..1010523)
                 /note="glycosyl transferase group 1"
                 /gene="BMETHMGA300858c"
CDS              complement(1010537..1011688)
                 /note="glycosyl transferase group 1"
                 /gene="BMETHMGA300859c"
CDS              complement(1013602..1014876)
                 /note="hypothetical protein"
                 /gene="BMETHMGA300860c"
CDS              complement(1014873..1016060)
                 /note="glycosyl transferase group 1"
                 /gene="BMETHMGA300861c"
CDS              complement(1016183..1017163)
                 /note="glycosyl transferase group 2"
                 /gene="BMETHMGA300862c"
CDS              complement(1017164..1018324)
                 /note="UDP-galactopyranose mutase"
                 /gene="BMETHMGA300863c"
CDS              complement(1018317..1019582)
                 /note="polysaccharide biosynthesis protein"
                 /gene="BMETHMGA300864c"
CDS              complement(1019721..1020407)
                 /note="sugar transferase"
                 /gene="BMETHMGA300865c"
CDS              complement(1020436..1020891)
                 /note="UDP-glucose-1-phosphate uridylyltransferase"
                 /gene="BMETHMGA300866c"
```

```
CDS             complement(1021676..1022446)
                /note="protein-tyrosine-phosphatase"
                /gene="BMETHMGA300867c"
CDS             complement(1022538..1023248)
                /note="capsular exopolysaccharide biosynthesis"
                /gene="BMETHMGA300868c"
CDS             complement(1023229..1023969)
                /note="lipopolysaccharide biosynthesis protein"
                /gene="BMETHMGA300869c"
CDS             complement(1024249..1025190)
                /note="cell envelope related function transcriptional
                regulator"
                /gene="BMETHMGA300870c"
CDS             complement(1025674..1026645)
                /note="UDP-glucose 4-epimerase"
                /gene="BMETHMGA300871c"
CDS             complement(1026712..1028274)
                /note="ATPase, AAA family"
                /gene="BMETHMGA300872c"
CDS             complement(1028271..1028876)
                /note="hypothetical protein"
                /gene="BMETHMGA300873c"
CDS             complement(1029028..1030728)
                /note="hypothetical protein"
                /gene="BMETHMGA300874c"
CDS             complement(1030736..1031218)
                /note="hypothetical protein"
                /gene="BMETHMGA300875c"
CDS             1031539..1032354
                /note="hypothetical protein"
                /gene="BMETHMGA300876"
CDS             1032409..1033350
                /note="hypothetical protein"
                /gene="BMETHMGA300877"
CDS             1033481..1034680
                /note="hypothetical protein"
                /gene="BMETHMGA300878"
CDS             complement(1034855..1035640)
                /note="5'-nucleotidase domain-containing protein"
                /gene="BMETHMGA300879c"
CDS             complement(1036138..1036713)
                /note="putative phage protein"
                /gene="BMETHMGA300880c"
CDS             complement(1037767..1039695)
                /note="modifier protein of major autolysin"
                /gene="BMETHMGA300881c"
CDS             complement(1039949..1040755)
                /note="hypothetical protein"
                /gene="BMETHMGA300882c"
CDS             complement(1040965..1041882)
                /note="AraC-type transcriptional regulator"
                /gene="BMETHMGA300883c"
CDS             1042077..1043240
```

```
                /note="major facilitator superfamily protein"
                /gene="BMETHMGA300884"
CDS             1043281..1044132
                /note="organophosphate/methylglyoxal/aldehyde
       reductase"
                /gene="BMETHMGA300885"
CDS             complement(1044364..1044828)
                /note="hypothetical protein"
                /gene="BMETHMGA300886c"
CDS             complement(1044899..1045336)
                /note="glyoxalase/bleomycin resistance
                protein/dioxygenase"
                /gene="BMETHMGA300887c"
CDS             1046665..1047915
                /note="hypothetical protein"
                /gene="BMETHMGA300888"
CDS             1047771..1048226
                /note="hypothetical protein"
                /gene="BMETHMGA300889"
CDS             complement(1048396..1048722)
                /note="Cupin 2 conserved barrel domain protein"
                /gene="BMETHMGA300890c"
CDS             complement(1048787..1049071)
                /note="nitric-oxide reductase"
                /gene="BMETHMGA300891c"
CDS             1049255..1049656
                /note="hypothetical protein"
                /gene="BMETHMGA300892"
CDS             1049875..1050408
                /note="nitroreductase"
                /gene="BMETHMGA300893"
CDS             complement(1050628..1050897)
                /note="exonuclease"
                /gene="BMETHMGA300894c"
CDS             complement(1051077..1052117)
                /note="copper amine oxidase domain-containing
       protein"
                /gene="BMETHMGA300895c"
CDS             complement(1052863..1055715)
                /note="endonuclease/exonuclease/phosphatase"
                /gene="BMETHMGA300896c"
CDS             complement(1055789..1059523)
                /note="2'3'-cyclic nucleotide
                2'-phosphodiesterase/3'-nucleotidase bifunctional
                periplasmic precursor protein"
                /gene="BMETHMGA300897c"
fasta_record    1060159..1113764
                /note="ConsensusfromContig23 Average coverage:
       608.65"
                /label=ConsensusfromContig23 Average coverage: 608.65
                /colour=10
CDS             1060513..1061451
                /note="Magnesium/cobalt transporter protein CorA"
```

```
                /gene="BMETHMGA300898"
CDS             complement(1062387..1062854)
                /note="SsrA-binding protein"
                /gene="BMETHMGA300899c"
CDS             complement(1062951..1065272)
                /note="ribonuclease R"
                /gene="BMETHMGA300900c"
CDS             complement(1065302..1066048)
                /note="thermostable carboxylesterase"
                /gene="BMETHMGA300901c"
CDS             complement(1066568..1067866)
                /note="phosphopyruvate hydratase"
                /gene="BMETHMGA300902c"
CDS             complement(1067888..1069423)
                /note="phosphoglyceromutase"
                /gene="BMETHMGA300903c"
CDS             complement(1069416..1070177)
                /note="trisephosphate isomrease"
                /gene="BMETHMGA300904c"
CDS             complement(1070207..1071391)
                /note="phosphoglycerate kinase"
                /gene="BMETHMGA300905c"
CDS             complement(1071543..1072550)
                /note="glyceraldehyde3-phosphate dehydrogenase"
                /gene="BMETHMGA300906c"
CDS             complement(1072598..1073662)
                /note="central glycolytic genes regulator"
                /gene="BMETHMGA300907c"
CDS             complement(1074027..1075364)
                /note="RNA polymerase sigma-54 factor"
                /gene="BMETHMGA300908c"
CDS             1075929..1076534
                /note="ATP-dependent Clp protease, proteolytic
       subunit"
                /gene="BMETHMGA300909"
CDS             complement(1076966..1077916)
                /note="hypothetical protein"
                /gene="BMETHMGA300910c"
CDS             complement(1077957..1078940)
                /note="hypothetical protein"
                /gene="BMETHMGA300911c"
CDS             complement(1078937..1079830)
                /note="hypothetical protein"
                /gene="BMETHMGA300912c"
CDS             complement(1079858..1080355)
                /note="NUDIX hydrolase"
                /gene="BMETHMGA300913c"
CDS             complement(1080702..1081598)
                /note="thioredoxin reductase"
                /gene="BMETHMGA300914c"
CDS             complement(1081772..1083268)
                /note="YvcD protein"
                /gene="BMETHMGA300915c"
```

```
CDS             complement(1083399..1084052)
                /note="bifunctional phosphoribosyl-AMP
                cyclohydrolase/phosphoribosyl-ATP pyrophosphatase"
                /gene="BMETHMGA300916c"
CDS             complement(1084049..1084807)
                /note="imidazole glycerol phosphate synthase subunit
       HisF"
                /gene="BMETHMGA300917c"
CDS             complement(1084807..1085538)
                /note="phosphoribosylformimino-5-aminoimidazole
                carboxamide ribotide isomerase"
                /gene="BMETHMGA300918c"
CDS             complement(1085622..1086260)
                /note="imidazole glycerol phosphate synthase,
       glutamine
                amidotransferase subunit"
                /gene="BMETHMGA300919c"
CDS             complement(1086326..1086913)
                /note="imidazoleglycerol-phosphate dehydratase"
                /gene="BMETHMGA300920c"
CDS             complement(1086955..1088034)
                /note="histidinol dehydrogenase"
                /gene="BMETHMGA300921c"
CDS             complement(1088245..1088892)
                /note="ATP phosphoribosyltransferase"
                /gene="BMETHMGA300922c"
CDS             complement(1088892..1090064)
                /note="ATp phosphoribosyltransferase regulatory
       subunit"
                /gene="BMETHMGA300923c"
CDS             complement(1090298..1090825)
                /note="putative acetyl-transferase"
                /gene="BMETHMGA300924c"
CDS             complement(1090829..1091473)
                /note="phosphoglycolate phosphatase"
                /gene="BMETHMGA300925c"
CDS             complement(1091463..1092413)
                /note="hypothetical protein YvoD"
                /gene="BMETHMGA300926c"
CDS             complement(1092461..1093282)
                /note="prolipoprotein diacylglyceryl transferase"
                /gene="BMETHMGA300927c"
CDS             complement(1093310..1094128)
                /note="HOr kinase/phosphorylase"
                /gene="BMETHMGA300928c"
CDS             1094497..1095708
                /note="N-acetylglucosamine-6-phosphate deacetylase"
                /gene="BMETHMGA300929"
CDS             1095686..1096411
                /note="N-acetylglucosamine-6-isomerase"
                /gene="BMETHMGA300930"
CDS             1096484..1097092
                /note="N-acetylmuramoyl-L-alanine amidase"
```

```
                     /gene="BMETHMGA300931"
     CDS             1097178..1098077
                     /note="acid phosphatase"
                     /gene="BMETHMGA300932"
     CDS             complement(1098161..1098517)
                     /note="hypothetical protein"
                     /gene="BMETHMGA300933c"
     CDS             complement(1098602..1099705)
                     /note="hypothetical protein"
                     /gene="BMETHMGA300934c"
     CDS             complement(1099884..1102757)
                     /note="excinuclease ABC, subunit A"
                     /gene="BMETHMGA300935c"
     CDS             complement(1102765..1104252)
                     /note="excinuclease ABC, subunit B"
                     /gene="BMETHMGA300936c"
     CDS             complement(1105524..1108778)
                     /note="acriflavin resistance protein"
                     /gene="BMETHMGA300937c"
     CDS             1108966..1109844
                     /note="transcriptional regulator, TetR family"
                     /gene="BMETHMGA300938"
     CDS             1110065..1110967
                     /note="hypothetical protein"
                     /gene="BMETHMGA300939"
     CDS             complement(1111145..1111813)
                     /note="hypothetical protein"
                     /gene="BMETHMGA300940c"
     CDS             complement(1111841..1112368)
                     /note="DinB family protein"
                     /gene="BMETHMGA300941c"
     CDS             complement(1112433..1113254)
                     /note="alpha/beta hydrolase"
                     /gene="BMETHMGA300942c"
     CDS             1113735..1114118
                     /note="hypothetical protein"
                     /gene="BMETHMGA300943"
     fasta_record    1113765..1157669
                     /note="ConsensusfromContig24 Average coverage:
            1,250.70"
                     /label=ConsensusfromContig24 Average coverage:
            1,250.70
                     /colour=11
     CDS             1115303..1115749
                     /note="DNA repair protein RadC"
                     /gene="BMETHMGA300944"
     CDS             1116983..1117594
                     /note="methyltransferase type II"
                     /gene="BMETHMGA300945"
     CDS             1118020..1118886
                     /note="hypothetical protein"
                     /gene="BMETHMGA300946"
     CDS             1119083..1120666
```

```
                     /note="multicopper oxidase"
                     /gene="BMETHMGA300947"
     CDS             1121261..1121968
                     /note="cytochrome c biogenesis protein"
                     /gene="BMETHMGA300948"
     CDS             complement(1122082..1122714)
                     /note="cytochrome c biogenesis protein"
                     /gene="BMETHMGA300949c"
     CDS             complement(1122627..1123247)
                     /note="cytochrome c biogenesis protein"
                     /gene="BMETHMGA300950c"
     CDS             complement(1123781..1124512)
                     /note="cytochrome c biogenesis protein transmembrane
                     protein"
                     /gene="BMETHMGA300951c"
     CDS             complement(1124636..1125610)
                     /note="glycosyl hydrolase"
                     /gene="BMETHMGA300952c"
     CDS             complement(1125723..1126292)
                     /note="hypothetical protein"
                     /gene="BMETHMGA300953c"
     CDS             complement(1126305..1126982)
                     /note="two-component response regulator"
                     /gene="BMETHMGA300954c"
     CDS             complement(1126979..1128148)
                     /note="two-component system histidine kinase"
                     /gene="BMETHMGA300955c"
     CDS             complement(1128787..1129089)
                     /note="transposase"
                     /gene="BMETHMGA300956c"
     CDS             complement(1129699..1130412)
                     /note="hypothetical protein"
                     /gene="BMETHMGA300957c"
     CDS             complement(1131465..1131905)
                     /note="hypothetical protein"
                     /gene="BMETHMGA300958c"
     CDS             complement(1131905..1133041)
                     /note="hypothetical protein"
                     /gene="BMETHMGA300959c"
     CDS             1133596..1135047
                     /note="replication protein"
                     /gene="BMETHMGA300960"
     CDS             1135811..1138354
                     /note="hypothetical protein"
                     /gene="BMETHMGA300961"
     CDS             1138366..1139256
                     /note="hypothetical protein"
                     /gene="BMETHMGA300962"
     CDS             1139632..1140609
                     /note="putative transposase"
                     /gene="BMETHMGA300964"
     CDS             1140916..1141359
                     /note="hypothetical protein"
```

```
                     /gene="BMETHMGA300965"
     CDS             1141337..1143919
                     /note="hypothetical protein"
                     /gene="BMETHMGA300966"
     CDS             1143916..1144359
                     /note="hypothetical protein"
                     /gene="BMETHMGA300967"
     CDS             1144356..1146497
                     /note="hypothetical protein"
                     /gene="BMETHMGA300968"
     CDS             1146620..1147147
                     /note="hypothetical protein"
                     /gene="BMETHMGA300969"
     CDS             1147134..1148324
                     /note="NLP/P60 family domain protein"
                     /gene="BMETHMGA300970"
     CDS             1150763..1151422
                     /note="phage related lysis protein"
                     /gene="BMETHMGA300973"
     CDS             1152014..1153063
                     /note="acylamide amidohydrolase"
                     /gene="BMETHMGA300974"
     CDS             1153906..1154808
                     /note="O-methyltransferase"
                     /gene="BMETHMGA300975"
     CDS             1154956..1155612
                     /note="cytochrome P450 family protein"
                     /gene="BMETHMGA300976"
     CDS             1155578..1155895
                     /note="cytochrome P450"
                     /gene="BMETHMGA300977"
     CDS             1156182..1156766
                     /note="transcriptional regulator, TetR family"
                     /gene="BMETHMGA300978"
     fasta_record    1157670..1237495
                     /note="ConsensusfromContig25 Average coverage:
            628.34"
                     /label=ConsensusfromContig25 Average coverage: 628.34
                     /colour=10
     CDS             1157889..1158365
                     /note="GntR family transcriptional regulator"
                     /gene="BMETHMGA300980"
     CDS             1158369..1159475
                     /note="hypothetical protein"
                     /gene="BMETHMGA300981"
     CDS             1159732..1160646
                     /note="arginase"
                     /gene="BMETHMGA300982"
     CDS             1161106..1161669
                     /note="RNA polymerase sigma-70 factor"
                     /gene="BMETHMGA300983"
     CDS             1161716..1162312
                     /note="anti-sigma factor RsiW"
```

```
                /gene="BMETHMGA300984"
CDS             1162714..1163544
                /note="hypothetical protein"
                /gene="BMETHMGA300985"
CDS             1163549..1164799
                /note="hypothetical protein YbbR"
                /gene="BMETHMGA300986"
CDS             1164836..1166188
                /note="phosphoglucosamine mutase"
                /gene="BMETHMGA300987"
CDS             1166684..1168486
                /note="fructose-6-phosphate amidotransferase"
                /gene="BMETHMGA300988"
CDS             1168922..1170322
                /note="amino acid transporter"
                /gene="BMETHMGA300989"
CDS             1170881..1171528
                /note="hypothetical protein"
                /gene="BMETHMGA300990"
CDS             1172255..1172629
                /note="3-demethylubiquinone-9 3-methyltransferase"
                /gene="BMETHMGA300991"
CDS             1172993..1174723
                /note="ABC transporter"
                /gene="BMETHMGA300992"
CDS             1174720..1176585
                /note="ABC transporter"
                /gene="BMETHMGA300993"
CDS             1177400..1178098
                /note="alpha/beta hydrolase fold protein"
                /gene="BMETHMGA300995"
CDS             1178385..1179848
                /note="PTS system, mannitol-specific IIC subunit"
                /gene="BMETHMGA300996"
CDS             1180082..1182205
                /note="PTS modulated transcriptional regulator"
                /gene="BMETHMGA300997"
CDS             1182207..1182644
                /note="PTS system mannitol-specific transporter
       subunit
                IIA"
                /gene="BMETHMGA300998"
CDS             1182646..1183806
                /note="mannitol dehydrogenase"
                /gene="BMETHMGA300999"
CDS             1184554..1184994
                /note="fosfomycin resistance protein"
                /gene="BMETHMGA301000"
CDS             1185302..1186228
                /note="transporter, drug metabolite exporter"
                /gene="BMETHMGA301001"
CDS             1186482..1187111
                /note="thiamine-phosphate diphosphorylase"
```

```
                     /gene="BMETHMGA301002"
     CDS             1187077..1188225
                     /note="glycine oxidase ThiO"
                     /gene="BMETHMGA301003"
     CDS             1188435..1189199
                     /note="thiazole synthase"
                     /gene="BMETHMGA301004"
     CDS             1189199..1190218
                     /note="thiamine/molybdopterin biosynthesis ThiF"
                     /gene="BMETHMGA301005"
     CDS             1190568..1191860
                     /note="gluconokinase"
                     /gene="BMETHMGA301006"
     CDS             1192411..1192878
                     /note="hypothetical protein"
                     /gene="BMETHMGA301007"
     CDS             complement(1192950..1194464)
                     /note="glycine betaine/Choline ABC transporter"
                     /gene="BMETHMGA301008c"
     CDS             complement(1194457..1195425)
                     /note="glycine betaine ABC transporter"
                     /gene="BMETHMGA301009c"
     CDS             1195750..1196568
                     /note="hypothetical protein"
                     /gene="BMETHMGA301010"
     CDS             complement(1196672..1198033)
                     /note="subtilisin-type proteinase"
                     /gene="BMETHMGA301011c"
     CDS             complement(1198164..1199588)
                     /note="amino acid permease"
                     /gene="BMETHMGA301012c"
     CDS             1200145..1200807
                     /note="hypothetical protein"
                     /gene="BMETHMGA301013"
     CDS             complement(1200860..1202218)
                     /note="Mg2+ transporter"
                     /gene="BMETHMGA301014c"
     CDS             1202734..1203390
                     /note="hypothetical protein"
                     /gene="BMETHMGA301015"
     CDS             1203837..1204814
                     /note="ferric ion ABC transporter"
                     /gene="BMETHMGA301016"
     CDS             1204969..1205835
                     /note="ferric ion ABC transporter, permease"
                     /gene="BMETHMGA301017"
     CDS             1205837..1207306
                     /note="iron(III) dicitrate transport system"
                     /gene="BMETHMGA301018"
     CDS             1207303..1208268
                     /note="cobalamin biosynthesis protein"
                     /gene="BMETHMGA301019"
     CDS             1208291..1209319
```

```
                     /note="aminotransferase"
                     /gene="BMETHMGA301020"
     CDS             1209301..1209909
                     /note="cobalbumin biosynthesis protein"
                     /gene="BMETHMGA301021"
     CDS             1209875..1211386
                     /note="cobyric acid synthase"
                     /gene="BMETHMGA301022"
     CDS             1211383..1212168
                     /note="cabalamin-5-phosphate synthase"
                     /gene="BMETHMGA301023"
     CDS             1212063..1212752
                     /note="alpha-ribazole-5'-phosphate phosphatase"
                     /gene="BMETHMGA301024"
     CDS             1212824..1213144
                     /note="cobU rlated cobalamine biosynthesis protein"
                     /gene="BMETHMGA301025"
     CDS             1213119..1213742
                     /note="ATP:cob(I) alamin adenosyltransferase"
                     /gene="BMETHMGA301026"
     CDS             1213742..1214236
                     /note="hypothetical protein"
                     /gene="BMETHMGA301027"
     CDS             1214233..1214949
                     /note="hypothetical protein"
                     /gene="BMETHMGA301028"
     CDS             1215014..1215646
                     /note="phosphoglycerate mutase"
                     /gene="BMETHMGA301029"
     CDS             1215883..1216332
                     /note="hypothetical protein"
                     /gene="BMETHMGA301030"
     CDS             1216799..1217881
                     /note="D-alanylalanine synthetase"
                     /gene="BMETHMGA301031"
     CDS             1217949..1219334
                     /note="UDP-N-acetylmuramoylalanyl-D-glutamyl-2,6-
          diaminopimelate"
                     /gene="BMETHMGA301032"
     CDS             1219468..1220199
                     /note="carboxylesterase"
                     /gene="BMETHMGA301033"
     CDS             1220449..1221870
                     /note="ATP-dependent RNA helicase"
                     /gene="BMETHMGA301034"
     CDS             1222063..1222542
                     /note="hypothetical protein"
                     /gene="BMETHMGA301035"
     CDS             1222529..1223971
                     /note="hypothetical protein"
                     /gene="BMETHMGA301036"
     CDS             complement(1224005..1224742)
                     /note="hypothetical protein"
```

```
                /gene="BMETHMGA301037c"
CDS             1225254..1226294
                /note="hypothetical protein"
                /gene="BMETHMGA301038"
CDS             1226680..1227735
                /note="alanine racemase"
                /gene="BMETHMGA301039"
CDS             1228459..1228809
                /note="endoribonuclease EndoA"
                /gene="BMETHMGA301040"
CDS             1229148..1229969
                /note="positive regulator of RsbR"
                /gene="BMETHMGA301041"
CDS             1229966..1230322
                /note="antagonist of RsbT"
                /gene="BMETHMGA301042"
CDS             1230327..1230728
                /note="RsbT protein"
                /gene="BMETHMGA301043"
CDS             1230739..1231749
                /note="phosphoserine phosphatase RsbU"
                /gene="BMETHMGA301044"
CDS             1231806..1232135
                /note="anti-anti-sigma factor (antagonist of RsbW)"
                /gene="BMETHMGA301045"
CDS             1232135..1232608
                /note="serine-protein kinase RsbW"
                /gene="BMETHMGA301046"
CDS             1232586..1233380
                /note="sigma factor B"
                /gene="BMETHMGA301047"
CDS             1233377..1233976
                /note="phosphoserine phosphatase rsbX"
                /gene="BMETHMGA301048"
CDS             1234040..1236268
                /note="transcription accessory protein"
                /gene="BMETHMGA301049"
CDS             1236444..1236968
                /note="hypothetical protein"
                /gene="BMETHMGA301050"
fasta_record    1237496..1461324
                /note="ConsensusfromContig26 Average coverage:
       530.24"
                /label=ConsensusfromContig26 Average coverage: 530.24
                /colour=11
CDS             1237739..1238140
                /note="hypothetical protein"
                /gene="BMETHMGA301051"
CDS             1238615..1239055
                /note="hypothetical protein"
                /gene="BMETHMGA301052"
CDS             1239171..1239560
                /note="hypothetical protein"
```

```
                     /gene="BMETHMGA301053"
     CDS             1239593..1239868
                     /note="hypothetical protein"
                     /gene="BMETHMGA301054"
     CDS             complement(1239900..1240196)
                     /note="hypothetical protein"
                     /gene="BMETHMGA301055c"
     CDS             1240485..1241909
                     /note="hypothetical protein"
                     /gene="BMETHMGA301056"
     CDS             1242230..1242790
                     /note="methyltransferase"
                     /gene="BMETHMGA301057"
     CDS             1242803..1243285
                     /note="phosphopantetheine adenylyltransferase"
                     /gene="BMETHMGA301058"
     CDS             complement(1243322..1244485)
                     /note="sporulation integral membrane protein"
                     /gene="BMETHMGA301059c"
     CDS             1244868..1245536
                     /note="patatin"
                     /gene="BMETHMGA301060"
     CDS             1245533..1246558
                     /note="PDZ/DHR/GLGF domain containing protein"
                     /gene="BMETHMGA301061"
     CDS             complement(1246619..1247857)
                     /note="hypothetical protein"
                     /gene="BMETHMGA301062c"
     CDS             1248139..1248546
                     /note="hypothetical protein"
                     /gene="BMETHMGA301063"
     CDS             1248941..1249702
                     /note="enoyl-CoA hydratase/isomerase"
                     /gene="BMETHMGA301064"
     CDS             1249760..1250422
                     /note="transcriptional factor, RsfA family"
                     /gene="BMETHMGA301065"
     CDS             1250584..1251975
                     /note="acetyl-CoA carboxylase"
                     /gene="BMETHMGA301066"
     CDS             1252222..1253763
                     /note="propionyl-CoA carboxylase"
                     /gene="BMETHMGA301067"
     CDS             1253913..1254815
                     /note="2-dehydropantoate 2-reductase"
                     /gene="BMETHMGA301068"
     CDS             1254817..1255212
                     /note="hypothetical protein"
                     /gene="BMETHMGA301069"
     CDS             1255311..1256945
                     /note="hypothetical protein"
                     /gene="BMETHMGA301070"
     CDS             1257164..1257595
```

```
                     /note="cell division protein MraZ"
                     /gene="BMETHMGA301071"
     CDS             1257656..1258588
                     /note="S-adenosyl-methyltransferase MraW"
                     /gene="BMETHMGA301072"
     CDS             1258621..1258974
                     /note="cell division protein FtsL"
                     /gene="BMETHMGA301073"
     CDS             1258974..1261184
                     /note="peptidoglycan glycosyltransferase"
                     /gene="BMETHMGA301074"
     CDS             1261285..1263201
                     /note="stage V sporulation protein D"
                     /gene="BMETHMGA301075"
     CDS             1263457..1264917
                     /note="UDP-N-acetylmuramoylalanyl-D-glutamate-2"
                     /gene="BMETHMGA301076"
     CDS             1265234..1266175
                     /note="phospho-N-acetylmuramoyl-pentapeptide-
          transferase"
                     /gene="BMETHMGA301077"
     CDS             1266437..1267528
                     /note="UDP-N-acetylmuramoylalanine-D-glutamate
          ligase"
                     /gene="BMETHMGA301078"
     CDS             1267627..1268727
                     /note="stage V sporulation protein E"
                     /gene="BMETHMGA301079"
     CDS             1269118..1270221
                     /note="N-acetylglucosaminyl transferase"
                     /gene="BMETHMGA301080"
     CDS             1270282..1271190
                     /note="UDP-N-acetylenolpyruvoylglucosamine reductase"
                     /gene="BMETHMGA301081"
     CDS             1271366..1272142
                     /note="cell-division initiation protein"
                     /gene="BMETHMGA301082"
     CDS             1272108..1272845
                     /note="hypothetical protein"
                     /gene="BMETHMGA301083"
     CDS             1272901..1273617
                     /note="hypothetical protein"
                     /gene="BMETHMGA301084"
     CDS             1273617..1273967
                     /note="small basic protein"
                     /gene="BMETHMGA301085"
     CDS             1274102..1275379
                     /note="cell divisiion protein FtsA"
                     /gene="BMETHMGA301086"
     CDS             1275421..1276560
                     /note="cell division protein FtsZ"
                     /gene="BMETHMGA301087"
     CDS             1276811..1277785
```

```
                /note="sporulation factor SpoIIGA"
                /gene="BMETHMGA301088"
CDS             1277824..1278543
                /note="RNA polymerase sigma-E factor"
                /gene="BMETHMGA301089"
CDS             1278770..1279480
                /note="RNA polymerase sigma-G"
                /gene="BMETHMGA301090"
CDS             1279926..1280738
                /note="hypothetical protein"
                /gene="BMETHMGA301091"
CDS             1280745..1281416
                /note="alanine racemase domain-containing protein"
                /gene="BMETHMGA301092"
CDS             1281435..1281872
                /note="hypothetical protein"
                /gene="BMETHMGA301093"
CDS             1282216..1282989
                /note="RNA-binding S4 domain protein"
                /gene="BMETHMGA301094"
CDS             1283184..1283594
                /note="cell division initiation protein"
                /gene="BMETHMGA301095"
CDS             1283962..1286736
                /note="isoleucyl-tRNA synthetase"
                /gene="BMETHMGA301096"
CDS             1286847..1287191
                /note="hypothetical protein"
                /gene="BMETHMGA301097"
CDS             1287226..1287765
                /note="signal peptidase II"
                /gene="BMETHMGA301098"
CDS             1287737..1288648
                /note="pseudouridine synthase"
                /gene="BMETHMGA301099"
CDS             1288865..1289410
                /note="bifunctional pyrimidine regulatory
       protein/uracil
                phosphoribosyltransferase"
                /gene="BMETHMGA301100"
CDS             1289544..1290857
                /note="uracil-xanthine permease"
                /gene="BMETHMGA301101"
CDS             1290989..1291972
                /note="aspartate carbamoyltransferase catalytic
       subunit"
                /gene="BMETHMGA301102"
CDS             1291935..1293221
                /note="dihydroorotase"
                /gene="BMETHMGA301103"
CDS             1293317..1294324
                /note="carbamoyl-phosphate synthase, small subunit"
                /gene="BMETHMGA301104"
```

```
CDS             1294303..1297515
                /note="carbamoyl-phosphate synthase, large subunit"
                /gene="BMETHMGA301105"
CDS             1297512..1298291
                /note="dihydroorotate dehydrogenase electron transfer
                subunit"
                /gene="BMETHMGA301106"
CDS             1298288..1299229
                /note="dihydroorotate dehydrogenase"
                /gene="BMETHMGA301107"
CDS             1299303..1299917
                /note="orotidine 5'-phosphate decarboxylase"
                /gene="BMETHMGA301108"
CDS             1299896..1300555
                /note="orotate phosphoribosyltransferase"
                /gene="BMETHMGA301109"
CDS             complement(1300652..1302376)
                /note="fibronectin-binding A domain protein"
                /gene="BMETHMGA301110c"
CDS             1302593..1305319
                /note="calcium-translocating P-type ATPase"
                /gene="BMETHMGA301111"
CDS             1305404..1306279
                /note="hypothetical protein"
                /gene="BMETHMGA301112"
CDS             1306612..1307229
                /note="guanylate kinase"
                /gene="BMETHMGA301113"
CDS             1307620..1308828
                /note="phosphopantothenoylcysteine
                decarboxylase/phosphopantothenate/cysteine ligase"
                /gene="BMETHMGA301114"
CDS             1308825..1311239
                /note="primosome assembly protein PriA"
                /gene="BMETHMGA301115"
CDS             1311403..1311786
                /note="peptide deformylase"
                /gene="BMETHMGA301116"
CDS             1311783..1312739
                /note="methionyl-tRNA formyltransferase"
                /gene="BMETHMGA301117"
CDS             1312729..1314069
                /note="sun protein"
                /gene="BMETHMGA301118"
CDS             1314116..1315165
                /note="ribosomal RNA large subunit methyltransferase"
                /gene="BMETHMGA301119"
CDS             1315113..1315928
                /note="serine/threonine phosphatase"
                /gene="BMETHMGA301120"
CDS             1315925..1317898
                /note="serine/threonine protein kinase"
                /gene="BMETHMGA301121"
```

```
CDS             1317913..1318794
                /note="ribosome-associated GTPase"
                /gene="BMETHMGA301122"
CDS             1318797..1319441
                /note="ribulose-phosphate 3-epimerase"
                /gene="BMETHMGA301123"
CDS             1319632..1320321
                /note="thiamine pyrophosphokinase"
                /gene="BMETHMGA301124"
CDS             1321166..1321528
                /note="hypothetical protein"
                /gene="BMETHMGA301125"
CDS             1321584..1323224
                /note="DAK phosphatase"
                /gene="BMETHMGA301126"
CDS             1323288..1323950
                /note="L-serine dehydratase, beta subunit"
                /gene="BMETHMGA301127"
CDS             1324018..1324896
                /note="L-serine dehydratase, alpha subunit"
                /gene="BMETHMGA301128"
CDS             1324760..1326946
                /note="ATP-dependent DNA helicase RecG"
                /gene="BMETHMGA301129"
CDS             1327241..1327813
                /note="transcriptional regulator of fatty acid
                biosynthesis"
                /gene="BMETHMGA301130"
CDS             1327746..1328828
                /note="glycerol-3-phosphate acetyltransferase"
                /gene="BMETHMGA301131"
CDS             1328818..1329786
                /note="malonyl CoA-acyl carrier protein transacylase"
                /gene="BMETHMGA301132"
CDS             1329786..1330529
                /note="beta-ketoacyl-acyl carrier protein reductase"
                /gene="BMETHMGA301133"
CDS             1330936..1331694
                /note="ribonuclease III"
                /gene="BMETHMGA301134"
CDS             1331899..1335288
                /note="chromosome segregation protein"
                /gene="BMETHMGA301135"
CDS             1335347..1336336
                /note="signal recognition particle-docking protein"
                /gene="BMETHMGA301136"
CDS             1336552..1336806
                /note="putative DNA-binding protein"
                /gene="BMETHMGA301137"
CDS             1336819..1338171
                /note="signal recognition particle protein"
                /gene="BMETHMGA301138"
CDS             1338282..1338554
```

```
                /note="30S ribosomal protein S16"
                /gene="BMETHMGA301139"
CDS             1338949..1339341
                /note="hypothetical protein"
                /gene="BMETHMGA301140"
CDS             1339350..1339868
                /note="16S rRNA-processing protein RimM"
                /gene="BMETHMGA301141"
CDS             1339865..1340608
                /note="tRNA (guanine-N(1)-) methyltransferase"
                /gene="BMETHMGA301142"
CDS             1340714..1341091
                /note="50S ribosomal protein L19"
                /gene="BMETHMGA301143"
CDS             1341382..1341948
                /note="signal peptidase I"
                /gene="BMETHMGA301144"
CDS             1341996..1342841
                /note="ribosomal biogenesis GTPase"
                /gene="BMETHMGA301145"
CDS             1342907..1343713
                /note="Ribonuclease H"
                /gene="BMETHMGA301146"
CDS             1343718..1345013
                /note="hypothetical protein"
                /gene="BMETHMGA301147"
CDS             1345010..1345300
                /note="flagellar biosynthesis protein"
                /gene="BMETHMGA301148"
CDS             1345505..1346662
                /note="succinyl-CoA synthetase subunit beta"
                /gene="BMETHMGA301149"
CDS             1346763..1347584
                /note="succinyl-CoA ligase subunit alpha"
                /gene="BMETHMGA301150"
CDS             1347670..1348530
                /note="DNA protecting protein DprA"
                /gene="BMETHMGA301151"
CDS             1348819..1350894
                /note="DNA topoisomerase I"
                /gene="BMETHMGA301152"
CDS             1351045..1352352
                /note="methylenetetrahydrofolate-tRNA-(uracil-5-)-
       methyltransferase"
                /gene="BMETHMGA301153"
CDS             1352378..1353316
                /note="tyrosine recombinase"
                /gene="BMETHMGA301154"
CDS             1353338..1353880
                /note="ATP-dependent protease peptidase subunit"
                /gene="BMETHMGA301155"
CDS             1353899..1355305
                /note="ATP-dependent protease ATP-binding subunit H"
```

```
                /gene="BMETHMGA301156"
CDS             1355426..1356205
                /note="transcriptional repressor CodY"
                /gene="BMETHMGA301157"
CDS             1356548..1357051
                /note="flagellar basal-body rod protein"
                /gene="BMETHMGA301158"
CDS             1357054..1357506
                /note="flagellar basal-body protein FlgC"
                /gene="BMETHMGA301159"
CDS             1357518..1357829
                /note="flagellar hook-basal body protein FliE"
                /gene="BMETHMGA301160"
CDS             1357900..1359486
                /note="Flagella M-ring protein FliF"
                /gene="BMETHMGA301161"
CDS             1359498..1360514
                /note="flagellar motor switch protein FliG"
                /gene="BMETHMGA301162"
CDS             1360498..1361271
                /note="flagellar assembly protein FliH"
                /gene="BMETHMGA301163"
CDS             1361268..1362587
                /note="flagellum-specific ATP synthase"
                /gene="BMETHMGA301164"
CDS             1362593..1363036
                /note="flagellar export protein FliJ"
                /gene="BMETHMGA301165"
CDS             1363045..1363659
                /note="hypothetical protein"
                /gene="BMETHMGA301166"
CDS             1363726..1364958
                /note="flagellar hook-length control protein"
                /gene="BMETHMGA301167"
CDS             1364968..1365393
                /note="flagellar basal body rod modification protein"
                /gene="BMETHMGA301168"
CDS             1365464..1366261
                /note="flagellar hook protein"
                /gene="BMETHMGA301169"
CDS             1366524..1366952
                /note="flagellar basal body-associated protein FliL"
                /gene="BMETHMGA301170"
CDS             1367000..1368019
                /note="flagellar motor switch protein FliM"
                /gene="BMETHMGA301171"
CDS             1367985..1369304
                /note="flagellar motor switch protein FliN"
                /gene="BMETHMGA301172"
CDS             1369318..1369686
                /note="two-component response regulator"
                /gene="BMETHMGA301174"
CDS             1369711..1370373
```

```
                     /note="flagellar biosynthesis protein FliZ"
                     /gene="BMETHMGA301175"
     CDS             1370370..1371035
                     /note="flagellar biosynthesis protein FliP"
                     /gene="BMETHMGA301176"
     CDS             1371050..1371319
                     /note="flagellar biosynthesis protein FliQ"
                     /gene="BMETHMGA301177"
     CDS             1371324..1372100
                     /note="flagellar biosynthesis protein FliR"
                     /gene="BMETHMGA301178"
     CDS             1372101..1373183
                     /note="flagellar biosynthesis protein FlhB"
                     /gene="BMETHMGA301179"
     CDS             1373240..1375276
                     /note="flagellar biosynthesis protein FlhA"
                     /gene="BMETHMGA301180"
     CDS             1375303..1376439
                     /note="flagellar biosynthesis regulator FlhF"
                     /gene="BMETHMGA301181"
     CDS             1376423..1377313
                     /note="antiactivator of flagellar biosynthesis FleN"
                     /gene="BMETHMGA301182"
     CDS             1377300..1378391
                     /note="chemotaxis-specific methylesterase"
                     /gene="BMETHMGA301183"
     CDS             1378384..1380477
                     /note="chemotaxis protein histidine kinase"
                     /gene="BMETHMGA301184"
     CDS             1380510..1380980
                     /note="chemotaxis signal transduction protein"
                     /gene="BMETHMGA301185"
     CDS             1380946..1381626
                     /note="chemotactic methyltransferase inhibitor"
                     /gene="BMETHMGA301186"
     CDS             1381619..1382116
                     /note="chemotaxis protein CheD"
                     /gene="BMETHMGA301187"
     CDS             1382282..1382950
                     /note="RNA polymerase sigma factor SigD"
                     /gene="BMETHMGA301188"
     CDS             1383015..1383650
                     /note="hypothetical protein"
                     /gene="BMETHMGA301189"
     CDS             1383799..1384512
                     /note="30S ribosomal protein S2"
                     /gene="BMETHMGA301190"
     CDS             1384614..1385495
                     /note="elongation factor Ts"
                     /gene="BMETHMGA301191"
     CDS             1385589..1386326
                     /note="uridylate kinase"
                     /gene="BMETHMGA301192"
```

```
CDS             1386326..1386883
                /note="ribosome recycling factor"
                /gene="BMETHMGA301193"
CDS             1387076..1387894
                /note="undecaprenyl diphosphate synthase"
                /gene="BMETHMGA301194"
CDS             1387960..1388748
                /note="phosphatidate cytidylyltransferase"
                /gene="BMETHMGA301195"
CDS             1388798..1389946
                /note="1-deoxy-D-xylulose 5-phosphate
       reductoisomerase"
                /gene="BMETHMGA301196"
CDS             1390293..1391354
                /note="membrane-associated zinc metalloprotease"
                /gene="BMETHMGA301197"
CDS             1391440..1393143
                /note="prolyl-tRNA synthetase"
                /gene="BMETHMGA301198"
CDS             1393341..1397663
                /note="DNA polymerase III, alpha subunit"
                /gene="BMETHMGA301199"
CDS             1397878..1398348
                /note="hypothetical protein"
                /gene="BMETHMGA301200"
CDS             1398406..1399521
                /note="transcription elongation protein NusA"
                /gene="BMETHMGA301201"
CDS             1399821..1400123
                /note="50S ribosomal protein"
                /gene="BMETHMGA301202"
CDS             1400143..1402371
                /note="translation initiation factor IF-2"
                /gene="BMETHMGA301203"
CDS             1402371..1402652
                /note="hypothetical protein"
                /gene="BMETHMGA301204"
CDS             1402708..1403079
                /note="ribosome-binding factor A"
                /gene="BMETHMGA301205"
CDS             1403175..1404053
                /note="tRNA pseudouridine synthase B"
                /gene="BMETHMGA301206"
CDS             1404151..1405038
                /note="bifunctional flavokinase/FAD synthetase"
                /gene="BMETHMGA301207"
CDS             1405706..1407820
                /note="polyribonucleotide nucleotidyltransferase"
                /gene="BMETHMGA301208"
CDS             1407998..1409008
                /note="sporulation protein"
                /gene="BMETHMGA301209"
CDS             1409257..1410342
```

```
                     /note="Zinc protease"
                     /gene="BMETHMGA301210"
     CDS             1410856..1411758
                     /note="dipicolinase synthase, A chain"
                     /gene="BMETHMGA301211"
     CDS             1411755..1412354
                     /note="dipicolinic acid synthetase, B subunit"
                     /gene="BMETHMGA301212"
     CDS             1412530..1413585
                     /note="aspartatei-semialdehyde dehydrogenase"
                     /gene="BMETHMGA301213"
     CDS             1413617..1414858
                     /note="aspartokinase I"
                     /gene="BMETHMGA301214"
     CDS             1414855..1415745
                     /note="dihydrodipicolinate synthase"
                     /gene="BMETHMGA301215"
     CDS             1416162..1417760
                     /note="zinc dependent hydrolase"
                     /gene="BMETHMGA301216"
     CDS             1417953..1418699
                     /note="translocation-enhancing protein"
                     /gene="BMETHMGA301217"
     CDS             1419081..1421417
                     /note="DNA translocase ftsK"
                     /gene="BMETHMGA301218"
     CDS             1421788..1422531
                     /note="transcriptional regulator, GntR family"
                     /gene="BMETHMGA301219"
     CDS             1422616..1423713
                     /note="ABC transporter"
                     /gene="BMETHMGA301220"
     CDS             1423934..1425469
                     /note="ABC transporter"
                     /gene="BMETHMGA301221"
     CDS             1425462..1426508
                     /note="ABC transporter permease"
                     /gene="BMETHMGA301222"
     CDS             1426518..1427468
                     /note="sugar ABC transporter"
                     /gene="BMETHMGA301223"
     CDS             1427696..1428979
                     /note="zinc dependent peptidase"
                     /gene="BMETHMGA301224"
     CDS             1429408..1430262
                     /note="peptidase M16 domain-containing protein"
                     /gene="BMETHMGA301226"
     CDS             1430355..1431077
                     /note="short-chain dehydrogenase/reductase"
                     /gene="BMETHMGA301227"
     CDS             1431616..1432407
                     /note="hypothetical protein"
                     /gene="BMETHMGA301228"
```

```
CDS             1432555..1433328
                /note="transcriptional regulator, XRE family"
                /gene="BMETHMGA301229"
CDS             1433378..1433971
                /note="phosphatidylglycerophosphate synthase"
                /gene="BMETHMGA301230"
CDS             1433996..1435246
                /note="competence damage-inducible protein A"
                /gene="BMETHMGA301231"
CDS             1435554..1436567
                /note="recombinase A"
                /gene="BMETHMGA301232"
CDS             1436849..1438408
                /note="phosphodiesterase"
                /gene="BMETHMGA301233"
CDS             1438494..1439291
                /note="metallophosphoesterase"
                /gene="BMETHMGA301234"
CDS             1440234..1441982
                /note="pyruvate synthase, alpha subunit"
                /gene="BMETHMGA301235"
CDS             1441969..1442835
                /note="ferrrodoxin oxidoreductase. beta subunit"
                /gene="BMETHMGA301236"
CDS             1443123..1444667
                /note="(dimethylallyl)adenosine tRNA
                methylthiotransferase"
                /gene="BMETHMGA301237"
CDS             1444723..1445100
                /note="hypothetical protein"
                /gene="BMETHMGA301238"
CDS             1445277..1445849
                /note="spore coat protein E"
                /gene="BMETHMGA301239"
CDS             1445975..1447030
                /note="hypothetical protein"
                /gene="BMETHMGA301240"
CDS             1447147..1449759
                /note="DNA mismatch repair protein MutS"
                /gene="BMETHMGA301241"
CDS             1449801..1451714
                /note="DNA mismatch repair protein MutL"
                /gene="BMETHMGA301242"
CDS             complement(1451875..1452888)
                /note="rod-shape determining protein"
                /gene="BMETHMGA301243c"
CDS             1453135..1454124
                /note="tRNA delta (2)-isopentenylpyrophosphate
                transferase"
                /gene="BMETHMGA301244"
CDS             1454655..1455596
                /note="stage V sporulation protein K"
                /gene="BMETHMGA301245"
```

```
CDS             1455958..1457133
                /note="GTP-binding protein HflX"
                /gene="BMETHMGA301246"
CDS             1457151..1458422
                /note="aluminum resistance protein"
                /gene="BMETHMGA301247"
CDS             1458683..1459063
                /note="transcriptional repressor of glutamine
       synthetase"
                /gene="BMETHMGA301248"
CDS             1459081..1460439
                /note="glutamine synthetase, type I"
                /gene="BMETHMGA301249"
fasta_record    1461325..1521976
                /note="ConsensusfromContig27 Average coverage:
       569.71"
                /label=ConsensusfromContig27 Average coverage: 569.71
                /colour=10
CDS             1461782..1463266
                /note="histidine ammonia-lyase"
                /gene="BMETHMGA301250"
CDS             1463715..1464071
                /note="hypothetical protein"
                /gene="BMETHMGA301251"
CDS             1464266..1465060
                /note="flagellar motor protein MotA"
                /gene="BMETHMGA301252"
CDS             1465053..1465811
                /note="flagellar motor protein MotB"
                /gene="BMETHMGA301253"
CDS             1466379..1467137
                /note="methionine aminopeptidase"
                /gene="BMETHMGA301254"
CDS             complement(1467215..1467469)
                /note="hypothetical protein"
                /gene="BMETHMGA301255c"
CDS             1467643..1467972
                /note="hypothetical protein"
                /gene="BMETHMGA301256"
CDS             1468021..1468842
                /note="Ribonuclease BN"
                /gene="BMETHMGA301257"
CDS             complement(1469212..1471131)
                /note="cadmium-transporting ATPase"
                /gene="BMETHMGA301258c"
CDS             complement(1471377..1472558)
                /note="major facilitator superfamily protein"
                /gene="BMETHMGA301259c"
CDS             1472808..1473866
                /note="calcium/proton exchanger"
                /gene="BMETHMGA301260"
CDS             1474239..1475093
                /note="hypothetical protein"
```

```
                     /gene="BMETHMGA301261"
     CDS             complement(1475309..1476430)
                     /note="putative Fe-S oxidoreductase"
                     /gene="BMETHMGA301262c"
     CDS             1476774..1478297
                     /note="fumarate hydratase, class I"
                     /gene="BMETHMGA301263"
     CDS             1478458..1479258
                     /note="delta-lactam-biosynthetic de-N-acetylase"
                     /gene="BMETHMGA301264"
     CDS             1479638..1480279
                     /note="DNA-3-methyladenine glycosylase II"
                     /gene="BMETHMGA301265"
     CDS             1480745..1481740
                     /note="23S rRNA (uracil-5-)-methyltransferase"
                     /gene="BMETHMGA301266"
     CDS             1483130..1484287
                     /note="NAD-dependent methanol dehydrogenase"
                     /gene="BMETHMGA301267"
     CDS             complement(1484420..1485334)
                     /note="nucleic acid binding protein"
                     /gene="BMETHMGA301268c"
     CDS             complement(1485345..1486289)
                     /note="permease family protein"
                     /gene="BMETHMGA301269c"
     CDS             1486555..1487514
                     /note="periplasmic solute binding protein"
                     /gene="BMETHMGA301270"
     CDS             1487710..1488117
                     /note="GCN5-related N-acetyltransferase"
                     /gene="BMETHMGA301271"
     CDS             complement(1488785..1489237)
                     /note="hypothetical protein"
                     /gene="BMETHMGA301272c"
     CDS             complement(1489487..1490293)
                     /note="methyltransferase type II"
                     /gene="BMETHMGA301273c"
     CDS             1490615..1492279
                     /note="urocanate hydratase"
                     /gene="BMETHMGA301274"
     CDS             1492349..1493650
                     /note="imidazolonepropionase"
                     /gene="BMETHMGA301275"
     CDS             1493623..1494621
                     /note="formimidoylglutamase"
                     /gene="BMETHMGA301276"
     CDS             complement(1495458..1496366)
                     /note="cell-division inhibitor"
                     /gene="BMETHMGA301277c"
     CDS             1496525..1497280
                     /note="RecA regulator RecX"
                     /gene="BMETHMGA301278"
     CDS             1497384..1497716
```

```
                     /note="hypothetical protein"
                     /gene="BMETHMGA301279"
     CDS             complement(1498644..1499600)
                     /note="membrane bound metal-dependent hydrolase"
                     /gene="BMETHMGA301281c"
     CDS             1499725..1500015
                     /note="hypothetical protein"
                     /gene="BMETHMGA301282"
     CDS             1500098..1501174
                     /note="A/G-specific adenine glycosylase"
                     /gene="BMETHMGA301283"
     CDS             1501544..1502293
                     /note="enoyl-(Acyl carrier protein) reductase"
                     /gene="BMETHMGA301284"
     CDS             1502828..1503079
                     /note="hypothetical protein"
                     /gene="BMETHMGA301285"
     CDS             1503395..1503925
                     /note="hypothetical protein"
                     /gene="BMETHMGA301286"
     CDS             1504229..1505785
                     /note="ABC transporter-like protein"
                     /gene="BMETHMGA301287"
     CDS             1505914..1507872
                     /note="alkaline phosphatase"
                     /gene="BMETHMGA301288"
     CDS             complement(1508033..1509106)
                     /note="hypothetical protein"
                     /gene="BMETHMGA301289c"
     CDS             complement(1509373..1513839)
                     /note="glutamate synthase, large subunit"
                     /gene="BMETHMGA301290c"
     CDS             complement(1514192..1515487)
                     /note="glutamate-1-semialdehyde aminotransferase"
                     /gene="BMETHMGA301291c"
     CDS             1515749..1516837
                     /note="ABC transporter"
                     /gene="BMETHMGA301292"
     CDS             1516782..1517639
                     /note="ABC transporter, permease"
                     /gene="BMETHMGA301293"
     CDS             1517643..1518428
                     /note="ABC transporter permease"
                     /gene="BMETHMGA301294"
     CDS             1518535..1518960
                     /note="hypothetical protein"
                     /gene="BMETHMGA301295"
     CDS             1519050..1519523
                     /note="peroxiredoxin"
                     /gene="BMETHMGA301296"
     CDS             1519824..1520267
                     /note="Fur family transcriptional regulator PerR"
                     /gene="BMETHMGA301297"
```

```
CDS             complement(1520360..1520704)
                /note="hypothetical protein"
                /gene="BMETHMGA301298c"
CDS             1520815..1521732
                /note="hypothetical protein"
                /gene="BMETHMGA301299"
fasta_record    1521977..1581027
                /note="ConsensusfromContig28 Average coverage:
       666.43"
                /label=ConsensusfromContig28 Average coverage: 666.43
                /colour=11
CDS             complement(1522223..1522453)
                /note="hypothetical protein"
                /gene="BMETHMGA301300c"
CDS             1522592..1523344
                /note="polysaccharide deacetylase"
                /gene="BMETHMGA301301"
CDS             complement(1523403..1524050)
                /note="integral membrane protein, activator of
       sporulation
                histidine kinase"
                /gene="BMETHMGA301302c"
CDS             1524187..1524807
                /note="spore germination protein GerD"
                /gene="BMETHMGA301303"
CDS             complement(1524918..1526054)
                /note="protein mrp salA"
                /gene="BMETHMGA301304c"
CDS             complement(1526174..1526896)
                /note="N-acetylmuramoyl-L-alanine amidase"
                /gene="BMETHMGA301305c"
CDS             complement(1526977..1527333)
                /note="hypothetical protein"
                /gene="BMETHMGA301306c"
CDS             complement(1527583..1527996)
                /note="30S ribosomal protein S9"
                /gene="BMETHMGA301307c"
CDS             complement(1527996..1528433)
                /note="50S ribosomal L13"
                /gene="BMETHMGA301308c"
CDS             complement(1528609..1529346)
                /note="tRNA pseudouridine synthase A"
                /gene="BMETHMGA301309c"
CDS             complement(1529360..1530157)
                /note="cobalt transport protein"
                /gene="BMETHMGA301310c"
CDS             complement(1530154..1531023)
                /note="ABC-type cobal transport system, ATPase"
                /gene="BMETHMGA301311c"
CDS             complement(1530999..1531838)
                /note="cobalt transporter ATP-binding subunit"
                /gene="BMETHMGA301312c"
CDS             complement(1532180..1532542)
```

```
                     /note="50S ribosomal protein L17"
                     /gene="BMETHMGA301313c"
     CDS             complement(1532619..1533563)
                     /note="DNA-directed RNA polymerase, alpha subunit"
                     /gene="BMETHMGA301314c"
     CDS             complement(1533737..1534126)
                     /note="30S ribosomal protein S11"
                     /gene="BMETHMGA301315c"
     CDS             complement(1534149..1534514)
                     /note="30S ribosomal protein S13"
                     /gene="BMETHMGA301316c"
     CDS             complement(1534896..1535222)
                     /note="hypothetical protein"
                     /gene="BMETHMGA301317c"
     CDS             complement(1535239..1535985)
                     /note="methionine aminopeptidase"
                     /gene="BMETHMGA301318c"
     CDS             complement(1535985..1536620)
                     /note="adenylate kinase"
                     /gene="BMETHMGA301319c"
     CDS             complement(1536694..1537989)
                     /note="preprotein translocase, SecY"
                     /gene="BMETHMGA301320c"
     CDS             complement(1537989..1538429)
                     /note="50S ribosomal L15"
                     /gene="BMETHMGA301321c"
     CDS             complement(1538657..1539157)
                     /note="30S ribosomal protein S5"
                     /gene="BMETHMGA301322c"
     CDS             complement(1539182..1539544)
                     /note="50S ribosomal protein L18"
                     /gene="BMETHMGA301323c"
     CDS             complement(1539577..1540113)
                     /note="50S ribosomal protein L6"
                     /gene="BMETHMGA301324c"
     CDS             complement(1540145..1540543)
                     /note="30S ribosomal protein S8"
                     /gene="BMETHMGA301325c"
     CDS             complement(1540785..1541324)
                     /note="50S ribosomal protein L5"
                     /gene="BMETHMGA301326c"
     CDS             complement(1541351..1541662)
                     /note="50S ribosomal protein L24"
                     /gene="BMETHMGA301327c"
     CDS             complement(1541699..1542067)
                     /note="50S ribosomal protein L14"
                     /gene="BMETHMGA301328c"
     CDS             complement(1542585..1543019)
                     /note="50S ribosomal protein L16"
                     /gene="BMETHMGA301329c"
     CDS             complement(1543022..1543678)
                     /note="30S ribosomal protein S3"
                     /gene="BMETHMGA301330c"
```

```
CDS             complement(1543682..1544044)
                /note="50S ribosomal protein L22"
                /gene="BMETHMGA301331c"
CDS             complement(1544044..1544322)
                /note="30S ribosomal protein S19"
                /gene="BMETHMGA301332c"
CDS             complement(1544383..1545213)
                /note="50S ribosomal protein L2"
                /gene="BMETHMGA301333c"
CDS             complement(1545245..1545529)
                /note="50S ribosomal protein L23"
                /gene="BMETHMGA301334c"
CDS             complement(1545529..1546152)
                /note="50S ribosomal protein L4"
                /gene="BMETHMGA301335c"
CDS             complement(1546174..1546833)
                /note="50S ribosomal protein L3"
                /gene="BMETHMGA301336c"
CDS             complement(1546845..1547153)
                /note="30S ribosomal protein S10"
                /gene="BMETHMGA301337c"
CDS             complement(1547457..1548644)
                /note="elongation factor Tu"
                /gene="BMETHMGA301338c"
CDS             complement(1548775..1550853)
                /note="elongation factor EF-2"
                /gene="BMETHMGA301339c"
CDS             complement(1550907..1551377)
                /note="30S ribosomal protein S7"
                /gene="BMETHMGA301340c"
CDS             complement(1551418..1551840)
                /note="30S ribosomal protein S12"
                /gene="BMETHMGA301341c"
CDS             complement(1551944..1552192)
                /note="50S ribosomal protein L7"
                /gene="BMETHMGA301342c"
CDS             complement(1552376..1555948)
                /note="DNA-directed RNA polymerase, beta subunit"
                /gene="BMETHMGA301343c"
CDS             complement(1556136..1559552)
                /note="DNA-directed RNA polymerase, beta subunit"
                /gene="BMETHMGA301344c"
CDS             complement(1559935..1560537)
                /note="ribosomal RNA methyltransferase"
                /gene="BMETHMGA301345c"
CDS             complement(1560668..1561033)
                /note="50S ribosomal protein L2/L12"
                /gene="BMETHMGA301346c"
CDS             complement(1561086..1561643)
                /note="50S ribosomal protein L10"
                /gene="BMETHMGA301347c"
CDS             complement(1561827..1562525)
                /note="50S ribosomal protein L1"
```

```
                /gene="BMETHMGA301348c"
CDS             complement(1562772..1563317)
                /note="50S ribosomal protein L11"
                /gene="BMETHMGA301349c"
CDS             complement(1563372..1563905)
                /note="transcription antitermination protein NusG"
                /gene="BMETHMGA301350c"
CDS             complement(1564554..1565252)
                /note="RNA polymerase sigma-70 factor"
                /gene="BMETHMGA301351c"
CDS             complement(1565278..1565784)
                /note="putative ribonuclease"
                /gene="BMETHMGA301352c"
CDS             complement(1565787..1566530)
                /note="tRNA/rRNA methyltransferase"
                /gene="BMETHMGA301353c"
CDS             complement(1566527..1566952)
                /note="ribonuclease III"
                /gene="BMETHMGA301354c"
CDS             complement(1566957..1568354)
                /note="cysteinyl-tRNA synthetase"
                /gene="BMETHMGA301355c"
CDS             complement(1568335..1569000)
                /note="serine acetyltransferase"
                /gene="BMETHMGA301356c"
CDS             complement(1569342..1570799)
                /note="glutamyl-tRNA synthetase"
                /gene="BMETHMGA301357c"
CDS             complement(1570878..1571357)
                /note="2-C-methyl-D-erythritol 2,4-
       cyclodiphosphatase"
                /gene="BMETHMGA301358c"
CDS             complement(1571388..1572086)
                /note="2-C-methyl-D-erythritol 4-phosphate
                cytidylyltransferase"
                /gene="BMETHMGA301359c"
CDS             complement(1572153..1573235)
                /note="putative membrane protein"
                /gene="BMETHMGA301360c"
CDS             complement(1573456..1574610)
                /note="DNA integrity scanning protein DisA"
                /gene="BMETHMGA301361c"
CDS             complement(1574532..1575911)
                /note="DNA repair protein RadA"
                /gene="BMETHMGA301362c"
CDS             complement(1576049..1578493)
                /note="class III stress response-related ATPase"
                /gene="BMETHMGA301363c"
CDS             complement(1578516..1579538)
                /note="ATP:guanido phosphotransferase"
                /gene="BMETHMGA301364c"
CDS             complement(1579583..1580017)
                /note="UvrB/UvrC protein"
```

```
                 /gene="BMETHMGA301365c"
CDS              complement(1580158..1580730)
                 /note="transcriptional regulator"
                 /gene="BMETHMGA301366c"
fasta_record     1581028..1654184
                 /note="ConsensusfromContig29 Average coverage:
       485.94"
                 /label=ConsensusfromContig29 Average coverage: 485.94
                 /colour=10
CDS              1582952..1583869
                 /note="putative cell wall-binding protein"
                 /gene="BMETHMGA301368"
CDS              complement(1584004..1584780)
                 /note="hypothetical protein"
                 /gene="BMETHMGA301369c"
CDS              1585696..1586058
                 /note="RDD domain-containing protein"
                 /gene="BMETHMGA301370"
CDS              1586387..1587028
                 /note="hypothetical protein"
                 /gene="BMETHMGA301371"
CDS              1587042..1587761
                 /note="hypothetical protein"
                 /gene="BMETHMGA301372"
CDS              1587784..1588491
                 /note="hypothetical protein"
                 /gene="BMETHMGA301373"
CDS              1588674..1588916
                 /note="hypothetical protein"
                 /gene="BMETHMGA301374"
CDS              1588994..1589500
                 /note="metal-dependent phosphohydrolase"
                 /gene="BMETHMGA301375"
CDS              complement(1590082..1591050)
                 /note="sporulation VanW like protein"
                 /gene="BMETHMGA301376c"
CDS              complement(1591137..1592540)
                 /note="transcriptional regulator, GntR family"
                 /gene="BMETHMGA301377c"
CDS              1592655..1593563
                 /note="EamA family transporter"
                 /gene="BMETHMGA301378"
CDS              1593614..1594447
                 /note="ABC transporter"
                 /gene="BMETHMGA301379"
CDS              complement(1594551..1595969)
                 /note="glycosyl transferase, family 2"
                 /gene="BMETHMGA301380c"
CDS              complement(1596133..1597344)
                 /note="multidrug resistance protein"
                 /gene="BMETHMGA301381c"
CDS              1597601..1598230
                 /note="biotin biosynthesis protein"
```

```
                /gene="BMETHMGA301382"
CDS             1598230..1598958
                /note="Dethiobiotin synthase"
                /gene="BMETHMGA301383"
CDS             1599143..1599568
                /note="OsmC family protein"
                /gene="BMETHMGA301384"
CDS             complement(1599657..1600655)
                /note="biotin synthase"
                /gene="BMETHMGA301385c"
CDS             1601063..1601992
                /note="glutaminase A"
                /gene="BMETHMGA301386"
CDS             complement(1602119..1602808)
                /note="hypothetical protein"
                /gene="BMETHMGA301387c"
CDS             1602925..1603236
                /note="hypothetical protein"
                /gene="BMETHMGA301388"
CDS             1603402..1604019
                /note="coat F domain protein"
                /gene="BMETHMGA301389"
CDS             complement(1604093..1605436)
                /note="amino-acid permease-associated protein"
                /gene="BMETHMGA301390c"
CDS             1605826..1607229
                /note="amino acid transporter"
                /gene="BMETHMGA301391"
CDS             1607705..1609069
                /note="adenosylmethionine-8-amino-7-oxononanoate
                transaminase"
                /gene="BMETHMGA301392"
CDS             1609405..1609851
                /note="transcriptional regulator, MarR family"
                /gene="BMETHMGA301393"
CDS             1609892..1611514
                /note="EmrB/QacA subfamily drug resistance
       transporter"
                /gene="BMETHMGA301394"
CDS             1611804..1612313
                /note="G/U mismatch-specific DNA glycosylase"
                /gene="BMETHMGA301395"
CDS             1612669..1613043
                /note="hypothetical protein"
                /gene="BMETHMGA301396"
CDS             1613525..1614916
                /note="amino acid permease-associated protein"
                /gene="BMETHMGA301397"
CDS             complement(1615144..1615755)
                /note="isochorismatase"
                /gene="BMETHMGA301398c"
CDS             1616001..1616531
                /note="hypothetical protein"
```

```
                /gene="BMETHMGA301399"
CDS             complement(1616982..1617566)
                /note="carbonic anhydrase"
                /gene="BMETHMGA301400c"
CDS             1618234..1618638
                /note="NADPH-dependent fmn reductase"
                /gene="BMETHMGA301401"
CDS             1618898..1619470
                /note="PfpI family protease"
                /gene="BMETHMGA301402"
CDS             1619635..1621026
                /note="Na+/H+ antiporter"
                /gene="BMETHMGA301403"
CDS             1621068..1621916
                /note="hypothetical protein"
                /gene="BMETHMGA301404"
CDS             complement(1622050..1622844)
                /note="short-chain dehydrogenase/reductase SDR"
                /gene="BMETHMGA301405c"
CDS             1623936..1625042
                /note="fatty acid desaturase"
                /gene="BMETHMGA301406"
CDS             1625152..1626291
                /note="two-component sensor histidine kinase, DesR"
                /gene="BMETHMGA301407"
CDS             1626288..1626887
                /note="two-component response regulator, DesR"
                /gene="BMETHMGA301408"
CDS             complement(1628233..1629453)
                /note="acetyl-CoA dehydrogenase"
                /gene="BMETHMGA301409c"
CDS             complement(1629552..1630631)
                /note="luciferase-like monooxygenase"
                /gene="BMETHMGA301410c"
CDS             1631269..1631727
                /note="riboflavin kinase"
                /gene="BMETHMGA301411"
CDS             1631645..1631995
                /note="riboflavin biosynthesis protein"
                /gene="BMETHMGA301412"
CDS             1633003..1633992
                /note="aliphatic sulfonates family ABC transporter"
                /gene="BMETHMGA301413"
CDS             1634011..1635159
                /note="alkanesulfonate monooxygenase"
                /gene="BMETHMGA301414"
CDS             1635159..1635923
                /note="alkanefulfonate transporter permease subunit"
                /gene="BMETHMGA301415"
CDS             1635939..1636706
                /note="Aliphatic sulfonates import ATP-binding
       protein"
                /gene="BMETHMGA301416"
```

```
CDS             1637063..1638133
                /note="branched-chain amino acid aminotransferase"
                /gene="BMETHMGA301417"
CDS             1638251..1638757
                /note="hypothetical protein"
                /gene="BMETHMGA301418"
CDS             1639061..1641247
                /note="glycerophosphodiester phosphodiesterase"
                /gene="BMETHMGA301419"
CDS             1641478..1642683
                /note="metabolite transport protein yceI"
                /gene="BMETHMGA301420"
CDS             1644084..1644899
                /note="hypothetical protein"
                /gene="BMETHMGA301421"
CDS             complement(1645121..1646341)
                /note="arabinose efflux protein"
                /gene="BMETHMGA301422c"
CDS             complement(1646364..1646807)
                /note="transcriptional regulatory protein"
                /gene="BMETHMGA301423c"
CDS             1647449..1647970
                /note="two-component transcriptional regulator"
                /gene="BMETHMGA301424"
CDS             1648014..1649309
                /note="histidine kinase"
                /gene="BMETHMGA301425"
CDS             1649467..1650840
                /note="hypothetical protein"
                /gene="BMETHMGA301426"
CDS             1651169..1651687
                /note="GCN5-related N-acetyltransferase"
                /gene="BMETHMGA301427"
CDS             1651914..1652252
                /note="hypothetical protein"
                /gene="BMETHMGA301428"
CDS             complement(1652746..1653948)
                /note="glycosyltransferase"
                /gene="BMETHMGA301429c"
fasta_record    1654185..1735406
                /note="ConsensusfromContig30 Average coverage:
       593.41"
                /label=ConsensusfromContig30 Average coverage: 593.41
                /colour=11
CDS             complement(1654715..1655386)
                /note="hypothetical protein"
                /gene="BMETHMGA301431c"
CDS             complement(1655412..1656092)
                /note="GCN5-related N-acetyltransferase"
                /gene="BMETHMGA301432c"
CDS             complement(1656079..1657029)
                /note="nitrilase/cyanide hydratase and apolipoprotein
                n-acetyltransferase"
```

```
                     /gene="BMETHMGA301433c"
CDS                  1657188..1658186
                     /note="Ion transport 2 domain-containing protein"
                     /gene="BMETHMGA301434"
CDS                  complement(1658183..1658590)
                     /note="hypothetical protein"
                     /gene="BMETHMGA301435c"
CDS                  complement(1658688..1660037)
                     /note="glucose-6-phosphate isomerase"
                     /gene="BMETHMGA301436c"
CDS                  complement(1660144..1661307)
                     /note="alcohol dehydrogenase (probably butanol)"
                     /gene="BMETHMGA301437c"
CDS                  1661310..1661708
                     /note="hypothetical protein"
                     /gene="BMETHMGA301438"
CDS                  complement(1661774..1662838)
                     /note="ornithine-oxo-acid transaminase"
                     /gene="BMETHMGA301439c"
CDS                  complement(1663261..1663665)
                     /note="general stress protein 13"
                     /gene="BMETHMGA301440c"
CDS                  complement(1663934..1665103)
                     /note="aminotransferase class I and II"
                     /gene="BMETHMGA301441c"
CDS                  complement(1665100..1665600)
                     /note="transcriptional regulator, AsnC family"
                     /gene="BMETHMGA301442c"
CDS                  1665765..1666592
                     /note="alpha/beta hydrolase"
                     /gene="BMETHMGA301443"
CDS                  complement(1666726..1667505)
                     /note="integral membrane protein, TerC-like"
                     /gene="BMETHMGA301444c"
CDS                  complement(1667834..1668397)
                     /note="superoxide dismutase, Cu-Zn"
                     /gene="BMETHMGA301445c"
CDS                  complement(1668838..1669371)
                     /note="peptidylprolyl isomerase"
                     /gene="BMETHMGA301446c"
CDS                  1669682..1670077
                     /note="kinase-associated lipoprotein"
                     /gene="BMETHMGA301447"
CDS                  complement(1670162..1670539)
                     /note="thioesterase superfamily protein"
                     /gene="BMETHMGA301448c"
CDS                  complement(1670615..1671220)
                     /note="alkaline phosphatase"
                     /gene="BMETHMGA301449c"
CDS                  1671450..1671869
                     /note="hypothetical protein"
                     /gene="BMETHMGA301450"
CDS                  complement(1672448..1673950)
```

```
                     /note="leucyl aminopeptidase"
                     /gene="BMETHMGA301451c"
     CDS             complement(1674011..1674391)
                     /note="hypothetical protein"
                     /gene="BMETHMGA301452c"
     CDS             1674535..1675014
                     /note="hypothetical protein"
                     /gene="BMETHMGA301453"
     CDS             1675024..1675806
                     /note="ABC transporter"
                     /gene="BMETHMGA301454"
     CDS             complement(1675936..1676649)
                     /note="hypothetical protein"
                     /gene="BMETHMGA301455c"
     CDS             complement(1676826..1677149)
                     /note="hypothetical protein"
                     /gene="BMETHMGA301456c"
     CDS             complement(1677245..1677736)
                     /note="NUDIX hydrolase"
                     /gene="BMETHMGA301457c"
     CDS             complement(1677815..1678981)
                     /note="NADH dehydrogenase"
                     /gene="BMETHMGA301458c"
     CDS             1679363..1680394
                     /note="thioredoxin reductase"
                     /gene="BMETHMGA301459"
     CDS             complement(1680485..1680847)
                     /note="iron-sulfur cluster assembly accessory
              protein"
                     /gene="BMETHMGA301460c"
     CDS             1681068..1681757
                     /note="hypothetical protein"
                     /gene="BMETHMGA301461"
     CDS             complement(1682035..1682274)
                     /note="UDP-N-acetylmuramoylalanine-D-glutamate
              liagase"
                     /gene="BMETHMGA301462c"
     CDS             1682683..1683813
                     /note="NADH dehydrogenase"
                     /gene="BMETHMGA301463"
     CDS             1684012..1684389
                     /note="hypothetical protein"
                     /gene="BMETHMGA301464"
     CDS             1684546..1685103
                     /note="inhibitor of cell-separation"
                     /gene="BMETHMGA301465"
     CDS             complement(1685300..1685680)
                     /note="putative sulfur oxido-reduction management
              enzyme,
                     YuzD"
                     /gene="BMETHMGA301466c"
     CDS             complement(1686307..1687224)
                     /note="homoserine kinase"
```

```
                   /gene="BMETHMGA301467c"
CDS                complement(1687221..1688282)
                   /note="threonine synthase"
                   /gene="BMETHMGA301468c"
CDS                complement(1688279..1689490)
                   /note="homoserine dehydrogenase I"
                   /gene="BMETHMGA301469c"
CDS                complement(1689792..1690889)
                   /note="spore coat protein S"
                   /gene="BMETHMGA301470c"
CDS                1691039..1691542
                   /note="phosphatidylglycerophosphatase A"
                   /gene="BMETHMGA301471"
CDS                complement(1691661..1692386)
                   /note="putative p-nitrophenyl phosphatase, YutF"
                   /gene="BMETHMGA301472c"
CDS                complement(1692521..1692958)
                   /note="hypothetical protein"
                   /gene="BMETHMGA301473c"
CDS                complement(1692985..1694301)
                   /note="diguanylate cyclase/phosphodiesterase"
                   /gene="BMETHMGA301474c"
CDS                1694763..1695074
                   /note="hypothetical protein"
                   /gene="BMETHMGA301475"
CDS                complement(1695208..1695483)
                   /note="hypothetical protein"
                   /gene="BMETHMGA301476c"
CDS                1695597..1696274
                   /note="sporulation lipoprotein YhcN/YlaJ-like
        protein"
                   /gene="BMETHMGA301477"
CDS                complement(1696436..1697356)
                   /note="lipoyl synthase"
                   /gene="BMETHMGA301478c"
CDS                1697647..1698540
                   /note="cell wall endopeptidase"
                   /gene="BMETHMGA301479"
CDS                1698820..1699998
                   /note="extracellular ligand-binding protein"
                   /gene="BMETHMGA301480"
CDS                1700085..1700963
                   /note="innner membrane translocator"
                   /gene="BMETHMGA301481"
CDS                1700978..1701952
                   /note="innner-membrane translocator"
                   /gene="BMETHMGA301482"
CDS                1701927..1702706
                   /note="ABC transporter"
                   /gene="BMETHMGA301483"
CDS                1702706..1703413
                   /note="ABC transporter related protein"
                   /gene="BMETHMGA301484"
```

```
CDS             complement(1703549..1704298)
                /note="sporulation protein"
                /gene="BMETHMGA301485c"
CDS             complement(1704384..1705469)
                /note="hypothetical protein"
                /gene="BMETHMGA301486c"
CDS             complement(1705725..1707116)
                /note="metallophosphoesterase"
                /gene="BMETHMGA301487c"
CDS             complement(1707203..1708054)
                /note="hypothetical protein"
                /gene="BMETHMGA301488c"
CDS             complement(1708252..1709649)
                /note="FeS assembly protein SufB"
                /gene="BMETHMGA301489c"
CDS             complement(1709698..1710141)
                /note="Fe-S cluster assembly protein NifU"
                /gene="BMETHMGA301490c"
CDS             complement(1710131..1711360)
                /note="cysteine desulhyrase"
                /gene="BMETHMGA301491c"
CDS             complement(1711360..1712670)
                /note="FeS assembly protein SufD"
                /gene="BMETHMGA301492c"
CDS             complement(1712688..1713473)
                /note="FeS assembly ATPase SufC"
                /gene="BMETHMGA301493c"
CDS             complement(1714127..1714969)
                /note="amino acid ABC transporter"
                /gene="BMETHMGA301494c"
CDS             complement(1714985..1715653)
                /note="ABC transporter"
                /gene="BMETHMGA301495c"
CDS             complement(1715646..1716524)
                /note="ABC transporter"
                /gene="BMETHMGA301496c"
CDS             complement(1717127..1718443)
                /note="O-acetylhomoserine sulfhydrylase"
                /gene="BMETHMGA301497c"
CDS             complement(1718884..1719195)
                /note="thioredoxin like protein"
                /gene="BMETHMGA301498c"
CDS             complement(1719197..1719637)
                /note="DNA primase"
                /gene="BMETHMGA301499c"
CDS             complement(1719932..1720315)
                /note="glycine cleavage system protein H"
                /gene="BMETHMGA301500c"
CDS             complement(1720413..1720772)
                /note="arsenate reductase"
                /gene="BMETHMGA301501c"
CDS             complement(1720928..1722727)
                /note="acetyl-CoA dehydrogenase"
```

```
                     /gene="BMETHMGA301502c"
     CDS             complement(1722772..1723947)
                     /note="acetyl-CoA acetyltransferase"
                     /gene="BMETHMGA301503c"
     CDS             complement(1723983..1726409)
                     /note="3-hyroxyacyl-CoA dehydrogenase/enoyl-CoA
                     hydratase/isomerase"
                     /gene="BMETHMGA301504c"
     CDS             complement(1726908..1727825)
                     /note="proline dehydrogenase"
                     /gene="BMETHMGA301506c"
     CDS             complement(1728009..1729424)
                     /note="glutamate synthase"
                     /gene="BMETHMGA301507c"
     CDS             1729568..1730314
                     /note="hypothetical protein"
                     /gene="BMETHMGA301508"
     CDS             complement(1730746..1731588)
                     /note="iron(III)-siderophore transporter"
                     /gene="BMETHMGA301509c"
     CDS             complement(1731623..1732681)
                     /note="iron(III) dicitrate ABC transporter"
                     /gene="BMETHMGA301510c"
     CDS             complement(1732678..1733682)
                     /note="Iron compound ABC transporter, permease"
                     /gene="BMETHMGA301511c"
     CDS             1734028..1734873
                     /note="iron(III) dicitrate ABC transporter"
                     /gene="BMETHMGA301512"
     CDS             1734992..1735432
                     /note="spore germination protein"
                     /gene="BMETHMGA301513"
     fasta_record    1735407..1753459
                     /note="ConsensusfromContig31 Average coverage:
            621.90"
                     /label=ConsensusfromContig31 Average coverage: 621.90
                     /colour=10
     CDS             1735525..1735779
                     /note="cobyrinic acid ac-diamine synthase"
                     /gene="BMETHMGA301514"
     CDS             1736898..1737200
                     /note="biotin synthase"
                     /gene="BMETHMGA301515"
     CDS             complement(1737379..1738590)
                     /note="drug resistance transporter"
                     /gene="BMETHMGA301516c"
     CDS             complement(1738580..1740229)
                     /note="hypothetical protein"
                     /gene="BMETHMGA301517c"
     CDS             complement(1741237..1741683)
                     /note="nitrogen fixation protein"
                     /gene="BMETHMGA301518c"
     CDS             complement(1742412..1742816)
```

```
                     /note="hypothetical protein"
                     /gene="BMETHMGA301519c"
     CDS             complement(1742968..1743264)
                     /note="30S ribosomal protein S14"
                     /gene="BMETHMGA301520c"
     CDS             complement(1743385..1744383)
                     /note="cobalamin synthesis protein"
                     /gene="BMETHMGA301521c"
     CDS             1744503..1745123
                     /note="dCTP deaminase"
                     /gene="BMETHMGA301522"
     CDS             1745329..1746249
                     /note="GTP cyclohydrolase"
                     /gene="BMETHMGA301523"
     CDS             1746806..1747816
                     /note="hyprothetical protein"
                     /gene="BMETHMGA301524"
     CDS             1748423..1749220
                     /note="cobyrinic acid ac-diamide synthase"
                     /gene="BMETHMGA301525"
     CDS             1749778..1750953
                     /note="hypothetical protein"
                     /gene="BMETHMGA301526"
     CDS             1751046..1752230
                     /note="cobalamin synthesis protein"
                     /gene="BMETHMGA301527"
     fasta_record    1753460..1848310
                     /note="ConsensusfromContig32 Average coverage:
            492.65"
                     /label=ConsensusfromContig32 Average coverage: 492.65
                     /colour=11
     CDS             complement(1753565..1755157)
                     /note="vanadium-dependent haloperoxidase"
                     /gene="BMETHMGA301528c"
     CDS             1755527..1756132
                     /note="hypothetical protein"
                     /gene="BMETHMGA301529"
     CDS             1756319..1757209
                     /note="transcriptional regulator, RpiR family"
                     /gene="BMETHMGA301530"
     CDS             1757222..1758376
                     /note="amidohydrolase"
                     /gene="BMETHMGA301531"
     CDS             1758360..1759190
                     /note="GCN5-related N-acetyltransferase"
                     /gene="BMETHMGA301532"
     CDS             1759210..1760316
                     /note="N-aceylamino acid racemase"
                     /gene="BMETHMGA301533"
     CDS             1760870..1762288
                     /note="hypothetical protein"
                     /gene="BMETHMGA301534"
     CDS             1762377..1763111
```

```
                     /note="putative orotidine 5'-phosphate decarboxylase"
                     /gene="BMETHMGA301535"
     CDS             1763127..1763477
                     /note="histidine biosynthesis protein"
                     /gene="BMETHMGA301536"
     CDS             1763822..1764595
                     /note="Cof-like hydrolase"
                     /gene="BMETHMGA301537"
     CDS             1765071..1765544
                     /note="cytochrome c oxidase, subunit II"
                     /gene="BMETHMGA301538"
     CDS             1765561..1767237
                     /note="cytochrome c oxidase subunit I"
                     /gene="BMETHMGA301539"
     CDS             1767602..1768108
                     /note="nitroreductase"
                     /gene="BMETHMGA301540"
     CDS             1768489..1769337
                     /note="hypothetical protein"
                     /gene="BMETHMGA301541"
     CDS             1769341..1770639
                     /note="tat-translocated enzyme"
                     /gene="BMETHMGA301542"
     CDS             1770672..1772429
                     /note="iron permease FTR1"
                     /gene="BMETHMGA301543"
     CDS             1772554..1773174
                     /note="ribokinase"
                     /gene="BMETHMGA301544"
     CDS             1773174..1773464
                     /note="ribokinase"
                     /gene="BMETHMGA301545"
     CDS             complement(1773639..1774142)
                     /note="hypothetical protein"
                     /gene="BMETHMGA301546c"
     CDS             1774706..1775665
                     /note="membrane dipeptidase"
                     /gene="BMETHMGA301547"
     CDS             1775883..1776968
                     /note="spermidine/putrescine ABC transporter"
                     /gene="BMETHMGA301548"
     CDS             1776980..1777783
                     /note="spermidine/putrescine ABC transporter"
                     /gene="BMETHMGA301549"
     CDS             1777780..1778580
                     /note="spermidine/putrescine ABC transporter"
                     /gene="BMETHMGA301550"
     CDS             1778577..1779650
                     /note="spermidine/putrescine import ABC transporter"
                     /gene="BMETHMGA301551"
     CDS             1779762..1780532
                     /note="hypothetical protein"
                     /gene="BMETHMGA301552"
```

```
CDS             complement(1780610..1781602)
                /note="hypothetical protein"
                /gene="BMETHMGA301553c"
CDS             1782510..1783877
                /note="aspartokinase III"
                /gene="BMETHMGA301554"
CDS             1784342..1785436
                /note="glutamate 5-kinase"
                /gene="BMETHMGA301555"
CDS             1785453..1786700
                /note="gamma-glutamyl phosphate reductase"
                /gene="BMETHMGA301556"
CDS             1787069..1787893
                /note="DPPA"
                /gene="BMETHMGA301557"
CDS             1787908..1788840
                /note="Dipeptide ABC transporter"
                /gene="BMETHMGA301558"
CDS             1789002..1789793
                /note="dipeptide ABC transporter"
                /gene="BMETHMGA301559"
CDS             1789799..1790797
                /note="dipeptide ABC transporter"
                /gene="BMETHMGA301560"
CDS             1790877..1792511
                /note="dipeptide ABC transporter"
                /gene="BMETHMGA301561"
CDS             1792599..1793522
                /note="L, D-carboxypeptidase"
                /gene="BMETHMGA301562"
CDS             1793906..1794700
                /note="L-Ala-D/L-Glu epimerase"
                /gene="BMETHMGA301563"
CDS             1794660..1795616
                /note="polysugar degrading enzyme"
                /gene="BMETHMGA301564"
CDS             1795620..1796630
                /note="oligopeptide ABC transporter"
                /gene="BMETHMGA301565"
CDS             1797077..1797742
                /note="phage shock protein A"
                /gene="BMETHMGA301566"
CDS             1797800..1798204
                /note="hypothetical protein"
                /gene="BMETHMGA301567"
CDS             1798219..1798734
                /note="hypothetical protein"
                /gene="BMETHMGA301568"
CDS             1798746..1799531
                /note="hypothetical protein"
                /gene="BMETHMGA301569"
CDS             complement(1799700..1800233)
                /note="putative transporter"
```

```
                     /gene="BMETHMGA301570c"
     CDS             1800508..1801128
                     /note="NUDIX hydrolase"
                     /gene="BMETHMGA301571"
     CDS             1801567..1802415
                     /note="molybdenum ABC transporter"
                     /gene="BMETHMGA301572"
     CDS             complement(1802428..1803576)
                     /note="Coenzyme PQQ synthesis protein"
                     /gene="BMETHMGA301573c"
     CDS             1803816..1804433
                     /note="nitroreductase"
                     /gene="BMETHMGA301574"
     CDS             1804778..1805839
                     /note="ABC transporter"
                     /gene="BMETHMGA301575"
     CDS             1805829..1806494
                     /note="ABC transporter permease"
                     /gene="BMETHMGA301576"
     CDS             1806524..1807357
                     /note="NLPA lipoprotein"
                     /gene="BMETHMGA301577"
     CDS             complement(1807447..1808259)
                     /note="formate dehydrogenase accessory protein"
                     /gene="BMETHMGA301578c"
     CDS             1808481..1810829
                     /note="formate dehyrogenase, alpha subunit"
                     /gene="BMETHMGA301579"
     CDS             complement(1810926..1811489)
                     /note="FAD-dependent pyridine nucleotide-disulfide
                     oxidoreductase"
                     /gene="BMETHMGA301580c"
     CDS             1812766..1813353
                     /note="alkaline phosphatase-like protein"
                     /gene="BMETHMGA301581"
     CDS             complement(1813684..1814802)
                     /note="proton glutamate symport protein"
                     /gene="BMETHMGA301582c"
     CDS             1815624..1816523
                     /note="alanyl-tRNA synthetase"
                     /gene="BMETHMGA301583"
     CDS             1816577..1816819
                     /note="alanyl-tRNA synthetase"
                     /gene="BMETHMGA301584"
     CDS             complement(1816972..1818024)
                     /note="chromate transporter"
                     /gene="BMETHMGA301585c"
     CDS             1818487..1819386
                     /note="hypothetical protein"
                     /gene="BMETHMGA301586"
     CDS             1819613..1819999
                     /note="transcriptional regulator, GntR family"
                     /gene="BMETHMGA301587"
```

```
CDS             1819996..1820895
                /note="ABC transporter"
                /gene="BMETHMGA301588"
CDS             1820849..1822819
                /note="hypothetical protein"
                /gene="BMETHMGA301589"
CDS             1823467..1825578
                /note="peptidoglycan glycosyltransferase"
                /gene="BMETHMGA301590"
CDS             1826184..1826996
                /note="hypothetical protein"
                /gene="BMETHMGA301591"
CDS             1827612..1829912
                /note="5-
       methyltetrahydropteroyltriglutamate/homocysteine
                S-methyltransferase"
                /gene="BMETHMGA301592"
CDS             1830142..1830570
                /note="hypothetical protein"
                /gene="BMETHMGA301593"
CDS             1830651..1831163
                /note="hypothetical protein"
                /gene="BMETHMGA301594"
CDS             complement(1831248..1832741)
                /note="ATP-dependent metalloprotease"
                /gene="BMETHMGA301595c"
CDS             complement(1833220..1836675)
                /note="methionine synthase"
                /gene="BMETHMGA301596c"
CDS             complement(1836653..1838518)
                /note="methylenetetrahydrofolate reductase"
                /gene="BMETHMGA301597c"
CDS             complement(1838515..1839717)
                /note="Cys/met metabolism pyridoxal-phosphate-
       dependent
                protein"
                /gene="BMETHMGA301598c"
CDS             complement(1839714..1840832)
                /note="Cys/Met metabolism pyridoxal-phosphate-
       dependent
                enzyme"
                /gene="BMETHMGA301599c"
CDS             1841236..1842021
                /note="ABC transporter"
                /gene="BMETHMGA301600"
CDS             1842156..1843130
                /note="NLPA lipoprotein"
                /gene="BMETHMGA301601"
CDS             1843135..1843908
                /note="ABC transporter"
                /gene="BMETHMGA301602"
CDS             complement(1844071..1844523)
                /note="transcriptional regulator, HxlR family"
```

```
                     /gene="BMETHMGA301603c"
CDS                  1844734..1845267
                     /note="hypothetical protein"
                     /gene="BMETHMGA301604"
CDS                  1845503..1846582
                     /note="acyl-CoA dehydrogenase"
                     /gene="BMETHMGA301605"
CDS                  complement(1846871..1847335)
                     /note="hypothetical protein"
                     /gene="BMETHMGA301606c"
CDS                  complement(1847385..1848197)
                     /note="micrococcal nuclease"
                     /gene="BMETHMGA301607c"
fasta_record         1848311..1920272
                     /note="ConsensusfromContig33 Average coverage:
         552.58"
                     /label=ConsensusfromContig33 Average coverage: 552.58
                     /colour=10
CDS                  complement(1848389..1848730)
                     /note="hypothetical protein"
                     /gene="BMETHMGA301608c"
CDS                  complement(1848747..1849955)
                     /note="stage II sporulatio protein P"
                     /gene="BMETHMGA301609c"
CDS                  complement(1850025..1851149)
                     /note="germination protease"
                     /gene="BMETHMGA301610c"
CDS                  complement(1851666..1852721)
                     /note="DNA polymerae III, delta subunit"
                     /gene="BMETHMGA301611c"
CDS                  complement(1853373..1855682)
                     /note="DNA internalization-related competence
         protein"
                     /gene="BMETHMGA301612c"
CDS                  complement(1855679..1856194)
                     /note="ComEB protein"
                     /gene="BMETHMGA301613c"
CDS                  complement(1856293..1856883)
                     /note="ComEA protein"
                     /gene="BMETHMGA301614c"
CDS                  1857000..1857821
                     /note="ComER protein"
                     /gene="BMETHMGA301615"
CDS                  complement(1857941..1858735)
                     /note="methyltransferase, type 12"
                     /gene="BMETHMGA301616c"
CDS                  complement(1858687..1859043)
                     /note="iojap-like protein"
                     /gene="BMETHMGA301617c"
CDS                  complement(1859048..1859614)
                     /note="hydrolase"
                     /gene="BMETHMGA301618c"
CDS                  complement(1859604..1860095)
```

```
                /note="nicotinate nucleotide adenylyltransferase"
                /gene="BMETHMGA301619c"
CDS             complement(1860613..1861470)
                /note="shikimate 5-dehydrogenase"
                /gene="BMETHMGA301621c"
CDS             complement(1861484..1862590)
                /note="ribosome biogenesis GTPase"
                /gene="BMETHMGA301622c"
CDS             complement(1862594..1863109)
                /note="hypothetical protein"
                /gene="BMETHMGA301623c"
CDS             complement(1863769..1864575)
                /note="phosphatidylserine decarboxylase"
                /gene="BMETHMGA301624c"
CDS             1864685..1865224
                /note="CDP-diacylglycerol/serine
                O-phosphatidyltransferase"
                /gene="BMETHMGA301625"
CDS             complement(1865300..1866019)
                /note="RNA polymerase sigma-K factor"
                /gene="BMETHMGA301626c"
CDS             complement(1866836..1867750)
                /note="cystathionine beta-lyase"
                /gene="BMETHMGA301627c"
CDS             complement(1867973..1868896)
                /note="cysteine synthase"
                /gene="BMETHMGA301628c"
CDS             complement(1868911..1869618)
                /note="5'-methylthioadenosine/S-adenosylhomocysteine
                nucleosidase"
                /gene="BMETHMGA301629c"
CDS             complement(1869631..1870269)
                /note="S-adenosylmethionine (SAM)-dependent
                methyltransferase"
                /gene="BMETHMGA301630c"
CDS             complement(1870694..1871368)
                /note="hypothetical protein"
                /gene="BMETHMGA301631c"
CDS             complement(1871464..1873236)
                /note="peptidoglycan glycosyltransferase"
                /gene="BMETHMGA301632c"
CDS             complement(1873407..1873859)
                /note="transcription elongantion factor GreA"
                /gene="BMETHMGA301633c"
CDS             complement(1874032..1874667)
                /note="uridine kinase"
                /gene="BMETHMGA301634c"
CDS             complement(1874672..1875940)
                /note="peptidase"
                /gene="BMETHMGA301635c"
CDS             complement(1875956..1876885)
                /note="peptidase"
                /gene="BMETHMGA301636c"
```

```
CDS             complement(1876891..1877460)
                /note="O-methyltransferase family 3"
                /gene="BMETHMGA301637c"
CDS             complement(1877769..1878902)
                /note="aminodeoxychorismate lyase"
                /gene="BMETHMGA301638c"
CDS             complement(1879130..1879420)
                /note="putative anti-sigma factor, YrzB"
                /gene="BMETHMGA301639c"
CDS             complement(1879433..1879849)
                /note="Holliday junction resolvase YqgF"
                /gene="BMETHMGA301640c"
CDS             complement(1880213..1882846)
                /note="alanyl-tRNA synthetase"
                /gene="BMETHMGA301641c"
CDS             complement(1883196..1884272)
                /note="tRNA-specific 3-thiouridylase MnmA"
                /gene="BMETHMGA301642c"
CDS             complement(1884843..1885319)
                /note="PRC-barrel domain-containing protein"
                /gene="BMETHMGA301643c"
CDS             complement(1885479..1887893)
                /note="helicase RecD/TraA family"
                /gene="BMETHMGA301644c"
CDS             complement(1887926..1888591)
                /note="putative tetratricopeptide repeat family
       protein"
                /gene="BMETHMGA301645c"
CDS             complement(1888791..1889867)
                /note="hypothetical protein"
                /gene="BMETHMGA301646c"
CDS             complement(1889937..1891043)
                /note="Cysteine desulfurase"
                /gene="BMETHMGA301647c"
CDS             complement(1891109..1891528)
                /note="transcriptional regulator of cysteine
       biosynthesis"
                /gene="BMETHMGA301648c"
CDS             complement(1891525..1892247)
                /note="hypothetical protein"
                /gene="BMETHMGA301649c"
CDS             complement(1892367..1893659)
                /note="replication-associated recombination protein
       A"
                /gene="BMETHMGA301650c"
CDS             complement(1894078..1895853)
                /note="apartyl-tRNA synthetase"
                /gene="BMETHMGA301651c"
CDS             complement(1895867..1897138)
                /note="histidyl-tRNA synthetase"
                /gene="BMETHMGA301652c"
CDS             1897963..1899735
                /note="hypothetical protein"
```

```
                     /gene="BMETHMGA301653"
     CDS             complement(1899889..1900329)
                     /note="D-tyrosyl-tRNA deacylase"
                     /gene="BMETHMGA301654c"
     CDS             complement(1900341..1902560)
                     /note="GTP pyrophosphokinase"
                     /gene="BMETHMGA301655c"
     CDS             complement(1902773..1903285)
                     /note="adenine phosphoribosyltransferase"
                     /gene="BMETHMGA301656c"
     CDS             complement(1903276..1905642)
                     /note="single-stranded-DNA-specific exonuclease RecJ"
                     /gene="BMETHMGA301657c"
     CDS             complement(1905788..1906687)
                     /note="cation efflux transporter
            (cadmium/zinc/cobalt)"
                     /gene="BMETHMGA301658c"
     CDS             complement(1906803..1909127)
                     /note="pretprotein translocase D and F"
                     /gene="BMETHMGA301659c"
     CDS             complement(1909273..1909569)
                     /note="post-transcriptional regulator"
                     /gene="BMETHMGA301660c"
     CDS             1909704..1911266
                     /note="stage V sporulation protein B"
                     /gene="BMETHMGA301661"
     CDS             complement(1911334..1911978)
                     /note="hypothetical protein"
                     /gene="BMETHMGA301662c"
     CDS             1912122..1912559
                     /note="hypothetical protein"
                     /gene="BMETHMGA301663"
     CDS             complement(1912749..1913369)
                     /note="transcriptional regulator TetR/AcrA family"
                     /gene="BMETHMGA301664c"
     CDS             complement(1913350..1914591)
                     /note="efflux transporter"
                     /gene="BMETHMGA301665c"
     CDS             complement(1915182..1916219)
                     /note="queuine tRNA-ribosyltransferase"
                     /gene="BMETHMGA301666c"
     CDS             complement(1916361..1917389)
                     /note="S-adenosylmethionine/tRNA-ribosyltransferase-
            isomerase"
                     /gene="BMETHMGA301667c"
     CDS             complement(1917618..1918631)
                     /note="Holliday jucntion DNA helicase B"
                     /gene="BMETHMGA301668c"
     CDS             complement(1918699..1919340)
                     /note="Holliday junction DNA helicase RuvA"
                     /gene="BMETHMGA301669c"
     CDS             complement(1919457..1920017)
                     /note="bypass-of-forespore protein C"
```

```
                /gene="BMETHMGA301670c"
CDS             complement(1920268..1920621)
                /note="spore germination protein"
                /gene="BMETHMGA301671c"
fasta_record    1920273..1941621
                /note="ConsensusfromContig34 Average coverage:
       1,225.94"
                /label=ConsensusfromContig34 Average coverage:
       1,225.94
                /colour=11
CDS             complement(1920902..1921186)
                /note="spore germination protein GerA"
                /gene="BMETHMGA301672c"
CDS             1921776..1922294
                /note="hypothetical protein"
                /gene="BMETHMGA301673"
CDS             1922695..1923273
                /note="hypothetical protein"
                /gene="BMETHMGA301674"
CDS             1924326..1925285
                /note="sodium/hyrogen exchanger"
                /gene="BMETHMGA301675"
CDS             complement(1925634..1926050)
                /note="peptidase U61 LD-carboxypeptidase A"
                /gene="BMETHMGA301676c"
CDS             complement(1926050..1926550)
                /note="peptidase U61 LD-carboxypeptidase A"
                /gene="BMETHMGA301677c"
CDS             1927254..1927880
                /note="nuclease"
                /gene="BMETHMGA301678"
CDS             1928740..1929195
                /note="hypothetical protein"
                /gene="BMETHMGA301679"
CDS             1929388..1929696
                /note="hypothetical protein"
                /gene="BMETHMGA301680"
CDS             complement(1930736..1931230)
                /note="hypothetical protein"
                /gene="BMETHMGA301681c"
CDS             complement(1932231..1933091)
                /note="BclI"
                /gene="BMETHMGA301683c"
CDS             1933461..1934336
                /note="M. BclI"
                /gene="BMETHMGA301684"
CDS             complement(1935858..1936976)
                /note="DNA helicase II/ATP-dependent helicase PcrA"
                /gene="BMETHMGA301686c"
CDS             complement(1936973..1938361)
                /note="ATP-dependent endonuclease of OLD family"
                /gene="BMETHMGA301687c"
CDS             1939908..1940330
```

```
                     /note="hypothetical protein"
                     /gene="BMETHMGA301688"
     CDS             complement(1940650..1940946)
                     /note="hypothetical protein"
                     /gene="BMETHMGA301689c"
     CDS             complement(1940930..1941616)
                     /note="cobyrinic acid ac-diamide synthase"
                     /gene="BMETHMGA301690c"
     fasta_record    1941622..1988931
                     /note="ConsensusfromContig35 Average coverage:
            579.88"
                     /label=ConsensusfromContig35 Average coverage: 579.88
                     /colour=10
     CDS             complement(1941696..1943282)
                     /note="phosphoenolpyruvate carboxykinase"
                     /gene="BMETHMGA301691c"
     CDS             1943488..1943871
                     /note="hyothetical protein"
                     /gene="BMETHMGA301692"
     CDS             1944401..1945678
                     /note="S-adenosylmethionine synthetase"
                     /gene="BMETHMGA301693"
     CDS             1946167..1948131
                     /note="asparagine synthase"
                     /gene="BMETHMGA301694"
     CDS             complement(1948128..1949000)
                     /note="hypothetical protein"
                     /gene="BMETHMGA301695c"
     CDS             complement(1949131..1949658)
                     /note="hypothetical protein"
                     /gene="BMETHMGA301696c"
     CDS             1949807..1950610
                     /note="alpha/beta hydrolase fold protein"
                     /gene="BMETHMGA301697"
     CDS             1950654..1951706
                     /note="hypothetical protein"
                     /gene="BMETHMGA301698"
     CDS             complement(1951769..1952341)
                     /note="SAM-dependent methyltransferase"
                     /gene="BMETHMGA301699c"
     CDS             complement(1952345..1953058)
                     /note="putative Fe-S oxidoreductase"
                     /gene="BMETHMGA301700c"
     CDS             1953503..1953802
                     /note="hypothetical protein"
                     /gene="BMETHMGA301701"
     CDS             1954104..1955291
                     /note="putative efflux transporter"
                     /gene="BMETHMGA301702"
     CDS             1955796..1958219
                     /note="leucyl-tRNA synthetase"
                     /gene="BMETHMGA301703"
     CDS             1958337..1958864
```

```
                /note="hypothetical protein"
                /gene="BMETHMGA301704"
CDS             1958881..1959195
                /note="Rhodanase domain protein"
                /gene="BMETHMGA301705"
CDS             1959425..1960582
                /note="hypothetical protein"
                /gene="BMETHMGA301706"
CDS             complement(1960674..1962209)
                /note="SSS sodium solute transporter superfamily
       protein"
                /gene="BMETHMGA301707c"
CDS             complement(1962233..1962574)
                /note="hypothetical protein"
                /gene="BMETHMGA301708c"
CDS             complement(1963115..1964392)
                /note="hypothetical protein"
                /gene="BMETHMGA301709c"
CDS             1964652..1966310
                /note="polysaccharide biosynthesis protein"
                /gene="BMETHMGA301710"
CDS             1966688..1967437
                /note="16S pseudouridylate synthase"
                /gene="BMETHMGA301711"
CDS             complement(1967601..1967945)
                /note="regulatory protein DeoR"
                /gene="BMETHMGA301712c"
CDS             1968447..1969157
                /note="ABC transporter"
                /gene="BMETHMGA301713"
CDS             1969154..1969597
                /note="hypothetical protein"
                /gene="BMETHMGA301714"
CDS             1969610..1970317
                /note="hypothetical protein"
                /gene="BMETHMGA301715"
CDS             1970615..1971883
                /note="succinyl-diaminopimelate desuccinylase"
                /gene="BMETHMGA301716"
CDS             complement(1972004..1972936)
                /note="cysteine synthase A"
                /gene="BMETHMGA301717c"
CDS             1973042..1973710
                /note="2'-5' RNA ligase"
                /gene="BMETHMGA301718"
CDS             1973758..1974759
                /note="hypothetical protein"
                /gene="BMETHMGA301719"
CDS             1975368..1977086
                /note="pullulanase"
                /gene="BMETHMGA301720"
CDS             1977473..1978072
                /note="aminogylcoside phosphotransferase"
```

```
                /gene="BMETHMGA301721"
CDS             1978490..1979164
                /note="tRNA (guanine-M(7)-)-methyltransferase"
                /gene="BMETHMGA301722"
CDS             complement(1979265..1979582)
                /note="hypothetical protein"
                /gene="BMETHMGA301723c"
CDS             1979821..1980894
                /note="glutamyl aminopeptidase"
                /gene="BMETHMGA301724"
CDS             1981055..1981588
                /note="hypothetical protein"
                /gene="BMETHMGA301725"
CDS             1981680..1982042
                /note="thioredoxin"
                /gene="BMETHMGA301726"
CDS             1982274..1983101
                /note="hypothetical protein"
                /gene="BMETHMGA301727"
CDS             1983388..1985967
                /note="DNA segregation ATPas FtsK/SpoIIIE"
                /gene="BMETHMGA301728"
CDS             1986146..1987447
                /note="UDP-N-acetylmuramate-L-alanine ligase"
                /gene="BMETHMGA301729"
CDS             1987669..1988169
                /note="hypothetical protein ytxG"
                /gene="BMETHMGA301730"
CDS             1988190..1988753
                /note="hypothetical protein ytxH"
                /gene="BMETHMGA301731"
fasta_record    1988932..2007216
                /note="ConsensusfromContig36 Average coverage:
       531.62"
                /label=ConsensusfromContig36 Average coverage: 531.62
                /colour=11
CDS             complement(1989021..1989620)
                /note="manganase oxidation"
                /gene="BMETHMGA301732c"
CDS             complement(1989811..1991295)
                /note="aldehyde dehydrogenase"
                /gene="BMETHMGA301733c"
CDS             1991665..1992036
                /note="LgrA protein"
                /gene="BMETHMGA301734"
CDS             1992149..1992694
                /note="LgrB family protein"
                /gene="BMETHMGA301735"
CDS             1993842..1994726
                /note="carbon monoxide dehydrogenae, medium subunit"
                /gene="BMETHMGA301736"
CDS             1994723..1995238
                /note="carbon monoxide dehydrogenase, small subunit"
```

```
                     /gene="BMETHMGA301737"
CDS                  1995249..1997570
                     /note="carbon monoxide dehydrogenase, large subunit"
                     /gene="BMETHMGA301738"
CDS                  1997647..1998654
                     /note="hypothetical protein"
                     /gene="BMETHMGA301739"
CDS                  1998778..1999221
                     /note="carbon monoxide dehydrogenase, subunit G"
                     /gene="BMETHMGA301740"
CDS                  1999239..2000099
                     /note="ATPase"
                     /gene="BMETHMGA301741"
CDS                  2000027..2001211
                     /note="carbon monoxide dehydrogenase, CoxE accessory
                     protein"
                     /gene="BMETHMGA301742"
CDS                  2001456..2002022
                     /note="cytochrome c oxidase assembly factor CtaG"
                     /gene="BMETHMGA301743"
CDS                  2002332..2003252
                     /note="Xanthine dehydrogenase"
                     /gene="BMETHMGA301744"
CDS                  2003271..2003903
                     /note="4-diphosphocytidyl-2C-methyl-D-erythritol
        synthase"
                     /gene="BMETHMGA301745"
CDS                  2003967..2005214
                     /note="molybdopterin biosynthesis protein moeA"
                     /gene="BMETHMGA301746"
CDS                  2005738..2006772
                     /note="major facilitator superfamily protein"
                     /gene="BMETHMGA301747"
fasta_record         2007217..2025337
                     /note="ConsensusfromContig37 Average coverage:
        576.99"
                     /label=ConsensusfromContig37 Average coverage: 576.99
                     /colour=10
CDS                  2008671..2009717
                     /note="monogalactosyldiacylglycerol synthase"
                     /gene="BMETHMGA301748"
CDS                  complement(2010439..2011866)
                     /note="hypothetical protein"
                     /gene="BMETHMGA301749c"
CDS                  complement(2012481..2013377)
                     /note="YegS/BmrU family lipid kinase"
                     /gene="BMETHMGA301750c"
CDS                  complement(2013446..2014348)
                     /note="hypothetical protein"
                     /gene="BMETHMGA301751c"
CDS                  2014570..2015766
                     /note="alcohol dehydrogenase GroES domain-containing
                     protein"
```

```
                     /gene="BMETHMGA301752"
     CDS             2015805..2016893
                     /note="class II aldolase/adducin family protein"
                     /gene="BMETHMGA301753"
     CDS             2016960..2017487
                     /note="prolipoprotein diacylglyceryl transferase"
                     /gene="BMETHMGA301754"
     CDS             2019085..2019720
                     /note="FMN-dependent NADH-azoreductase"
                     /gene="BMETHMGA301755"
     CDS             complement(2020226..2020759)
                     /note="hypothetical protein"
                     /gene="BMETHMGA301756c"
     CDS             complement(2021593..2021943)
                     /note="transcriptional regulator, MarR family"
                     /gene="BMETHMGA301757c"
     CDS             2022793..2023494
                     /note="SapB protein"
                     /gene="BMETHMGA301758"
     CDS             2023606..2024511
                     /note="hypothetical protein"
                     /gene="BMETHMGA301759"
     fasta_record    2025338..2039717
                     /note="ConsensusfromContig38 Average coverage:
            522.10"
                     /label=ConsensusfromContig38 Average coverage: 522.10
                     /colour=11
     CDS             complement(2025969..2027312)
                     /note="amino acid permease-associated protein"
                     /gene="BMETHMGA301760c"
     CDS             2027964..2029388
                     /note="arginine/ornithine antiporter"
                     /gene="BMETHMGA301761"
     CDS             complement(2029561..2029818)
                     /note="hypothetical protein"
                     /gene="BMETHMGA301762c"
     CDS             2030104..2030772
                     /note="hypothetical protein"
                     /gene="BMETHMGA301763"
     CDS             2031073..2031804
                     /note="DSBA oxidoreductase"
                     /gene="BMETHMGA301764"
     CDS             2032132..2033046
                     /note="ribosylpyrimidine nucleoside"
                     /gene="BMETHMGA301765"
     CDS             2033960..2035924
                     /note="putative thiazole-containing bacteriocin
            maturation
                     protein"
                     /gene="BMETHMGA301766"
     CDS             2036227..2037867
                     /note="bacteriocin biosynthesis docking scaffold,
            SagD
```

```
                family domain protein"
                /gene="BMETHMGA301767"
CDS             2037890..2039467
                /note="NADH oxidase"
                /gene="BMETHMGA301768"
fasta_record    2039718..2104498
                /note="ConsensusfromContig39 Average coverage:
       495.64"
                /label=ConsensusfromContig39 Average coverage: 495.64
                /colour=10
CDS             2039886..2041976
                /note="transcriptional antiterminator BglG"
                /gene="BMETHMGA301769"
CDS             2042076..2042465
                /note="PTS IIA-like nitrogen-regulatory protein PtsN"
                /gene="BMETHMGA301770"
CDS             2042850..2044112
                /note="PTS system, galactitol-specific IIC subunit"
                /gene="BMETHMGA301771"
CDS             2044147..2045178
                /note="alcohol dehydrogenase, GroES"
                /gene="BMETHMGA301772"
CDS             2045489..2046553
                /note="alcohol dehydrogenase GroES/Sorbitol
       dehydrogenase"
                /gene="BMETHMGA301774"
CDS             2046861..2047907
                /note="translation initiation factor IF-2B, alpha
       subunit"
                /gene="BMETHMGA301775"
CDS             2048022..2048540
                /note="thiol:disulfide interchange protein tlpA"
                /gene="BMETHMGA301776"
CDS             2050093..2050533
                /note="thioesterase superfamily protein"
                /gene="BMETHMGA301778"
CDS             complement(2050950..2051249)
                /note="HesB/YadR/YfhF-family protein"
                /gene="BMETHMGA301779c"
CDS             complement(2051412..2052551)
                /note="capsule synthesis protein, CapA"
                /gene="BMETHMGA301780c"
CDS             2053382..2054908
                /note="carbohydrate kinase"
                /gene="BMETHMGA301781"
CDS             2055217..2057559
                /note="UvrD/REP helicase"
                /gene="BMETHMGA301782"
CDS             2057609..2059264
                /note="EmrB/QacA subfamily drug resistance
       transporter"
                /gene="BMETHMGA301783"
CDS             complement(2059471..2060559)
```

```
                     /note="ATP-binding region ATPase domain protein"
                     /gene="BMETHMGA301784c"
     CDS             complement(2060862..2061446)
                     /note="hypothetical protein"
                     /gene="BMETHMGA301785c"
     CDS             2061980..2062603
                     /note="hypothetical protein"
                     /gene="BMETHMGA301786"
     CDS             2062570..2063505
                     /note="amidase"
                     /gene="BMETHMGA301787"
     CDS             complement(2063814..2064326)
                     /note="DoxX family protein"
                     /gene="BMETHMGA301788c"
     CDS             2064844..2066262
                     /note="spore peptidoglycan hydrolase"
                     /gene="BMETHMGA301789"
     CDS             2066600..2067013
                     /note="CoA-binding protein"
                     /gene="BMETHMGA301790"
     CDS             2067297..2069279
                     /note="DNA topoisomerase IV, B subunit"
                     /gene="BMETHMGA301791"
     CDS             2069279..2071738
                     /note="DNA topoisomerase IV, A subunit"
                     /gene="BMETHMGA301792"
     CDS             2071939..2072688
                     /note="hypothetical protein"
                     /gene="BMETHMGA301793"
     CDS             2073490..2073996
                     /note="integrase/recombinase"
                     /gene="BMETHMGA301794"
     CDS             complement(2074220..2074780)
                     /note="Endonuclease, Uma2 family"
                     /gene="BMETHMGA301795c"
     CDS             complement(2075479..2076075)
                     /note="integrase/recombinase"
                     /gene="BMETHMGA301796c"
     CDS             complement(2076702..2078732)
                     /note="hypothetical protein"
                     /gene="BMETHMGA301797c"
     CDS             complement(2079069..2079860)
                     /note="hypothetical protein"
                     /gene="BMETHMGA301798c"
     CDS             complement(2079971..2080519)
                     /note="hypothetical protein"
                     /gene="BMETHMGA301799c"
     CDS             complement(2081375..2081797)
                     /note="NUDIX hydrolase"
                     /gene="BMETHMGA301800c"
     CDS             complement(2082786..2083850)
                     /note="hypothetical protein"
                     /gene="BMETHMGA301801c"
```

```
CDS             complement(2083847..2084425)
                /note="RNA polymerase, sigma-24 subunit, ECF
       subfamily"
                /gene="BMETHMGA301802c"
CDS             complement(2085112..2087586)
                /note="tetratricopeptide-like helical"
                /gene="BMETHMGA301803c"
CDS             complement(2090566..2091948)
                /note="peptidase M28"
                /gene="BMETHMGA301805c"
CDS             complement(2092018..2093118)
                /note="hypothetical protein"
                /gene="BMETHMGA301806c"
CDS             2094499..2094819
                /note="ArsR family transcriptional regulator"
                /gene="BMETHMGA301807"
CDS             2094841..2095479
                /note="MIP fmaily channel protein"
                /gene="BMETHMGA301808"
CDS             2095534..2095956
                /note="arsenate reductase"
                /gene="BMETHMGA301809"
CDS             complement(2096465..2096851)
                /note="hypothetical protein"
                /gene="BMETHMGA301810c"
CDS             complement(2097104..2097913)
                /note="sulfate ABC transporter ATP-binding protein
       CysA"
                /gene="BMETHMGA301811c"
CDS             complement(2098210..2099076)
                /note="Sulfate ABC transporter, inner membrane
       subunit"
                /gene="BMETHMGA301812c"
CDS             complement(2099089..2099985)
                /note="sulfate ABC transporter, inner membrane
       subunit
                CysT"
                /gene="BMETHMGA301813c"
CDS             complement(2100018..2101097)
                /note="sulfate ABC transporter, sulfate-binding
       protein"
                /gene="BMETHMGA301814c"
CDS             complement(2101279..2102058)
                /note="NAD-dependent deacetylase/silent informatin
                regulation protein Sir2"
                /gene="BMETHMGA301815c"
CDS             2102947..2103807
                /note="transcriptional regulator, AraC family"
                /gene="BMETHMGA301816"
fasta_record    2104499..2114205
                /note="ConsensusfromContig40 Average coverage:
       552.22"
                /label=ConsensusfromContig40 Average coverage: 552.22
```

```
                   /colour=11
CDS                2105913..2107256
                   /note="phenylacetate-CoA ligase"
                   /gene="BMETHMGA301819"
CDS                2107288..2108523
                   /note="phenylacetate-CoA ligase"
                   /gene="BMETHMGA301820"
CDS                2108609..2109436
                   /note="bifunctional 3-deoxy-7-phosphoheptulonate
                   synthase/chorismate mutase"
                   /gene="BMETHMGA301821"
CDS                2109494..2110090
                   /note="hypothetical protein"
                   /gene="BMETHMGA301822"
CDS                2112651..2113856
                   /note="methylthioribose kinase"
                   /gene="BMETHMGA301825"
fasta_record       2114206..2117937
                   /note="ConsensusfromContig41 Average coverage:
       613.49"
                   /label=ConsensusfromContig41 Average coverage: 613.49
                   /colour=10
CDS                2114414..2114845
                   /note="spore cortex-lytic enzyme"
                   /gene="BMETHMGA301826"
CDS                2116148..2116528
                   /note="nucleotidyltransferase"
                   /gene="BMETHMGA301827"
CDS                2116500..2116922
                   /note="hypothetical protein"
                   /gene="BMETHMGA301828"
fasta_record       2117938..2173770
                   /note="ConsensusfromContig42 Average coverage:
       639.05"
                   /label=ConsensusfromContig42 Average coverage: 639.05
                   /colour=11
CDS                complement(2118297..2119733)
                   /note="lysyl-tRNA synthetase"
                   /gene="BMETHMGA301831c"
CDS                complement(2119980..2120981)
                   /note="tRNA-dihydrouridine synthase"
                   /gene="BMETHMGA301832c"
CDS                complement(2121182..2121706)
                   /note="2-amino-4-hydroxy-6-
       dhyroxymethyldihydropteridine
                   pyrophosphokinase"
                   /gene="BMETHMGA301833c"
CDS                complement(2121710..2122069)
                   /note="dihydroneopterin aldolase"
                   /gene="BMETHMGA301834c"
CDS                complement(2122069..2122836)
                   /note="dihydropteroate synthase"
                   /gene="BMETHMGA301835c"
```

```
CDS             complement(2122915..2123775)
                /note="4-amino-4-deoxychorismate lyase"
                /gene="BMETHMGA301836c"
CDS             complement(2123781..2124389)
                /note="para-aminobenzoate synthase component II"
                /gene="BMETHMGA301837c"
CDS             complement(2124373..2125638)
                /note="para-aminobenzoate synthase, component I"
                /gene="BMETHMGA301838c"
CDS             complement(2126101..2127024)
                /note="cysteine synthase"
                /gene="BMETHMGA301839c"
CDS             complement(2127167..2128066)
                /note="putative protein secretion PrsA homolog "
                /gene="BMETHMGA301840c"
CDS             complement(2128076..2128954)
                /note="Hsp33-like chaperonin"
                /gene="BMETHMGA301841c"
CDS             complement(2129013..2129669)
                /note="transcriptional activator, Baf family"
                /gene="BMETHMGA301842c"
CDS             complement(2129972..2131954)
                /note="cell division protein FtsH"
                /gene="BMETHMGA301843c"
CDS             complement(2132085..2132657)
                /note="hypoxanthine phosphoribosyltransferase"
                /gene="BMETHMGA301844c"
CDS             complement(2132654..2134042)
                /note="hypoxanthine phosphoribosyltransferase"
                /gene="BMETHMGA301845c"
CDS             complement(2134232..2135215)
                /note="serine/threonine protein kinase"
                /gene="BMETHMGA301846c"
CDS             complement(2135184..2135954)
                /note="hypothetical protein"
                /gene="BMETHMGA301847c"
CDS             complement(2136096..2138201)
                /note="stage II sporulation protein E"
                /gene="BMETHMGA301848c"
CDS             complement(2139090..2139539)
                /note="RNA binding S1 domain protein"
                /gene="BMETHMGA301849c"
CDS             complement(2139623..2140027)
                /note="hypothetical protein"
                /gene="BMETHMGA301850c"
CDS             complement(2140034..2140666)
                /note="spore cortex biosynthesis protein YabQ"
                /gene="BMETHMGA301851c"
CDS             complement(2140663..2140974)
                /note="sporulation protein YabP"
                /gene="BMETHMGA301852c"
CDS             complement(2141366..2142829)
                /note="MazG family protein"
```

```
                     /gene="BMETHMGA301853c"
     CDS             complement(2142842..2144449)
                     /note="polysaccharide biosynthesis protein"
                     /gene="BMETHMGA301854c"
     CDS             complement(2144604..2145140)
                     /note="stage V sporulation protein T"
                     /gene="BMETHMGA301855c"
     CDS             complement(2145418..2148948)
                     /note="transcription-repair coupling factor"
                     /gene="BMETHMGA301856c"
     CDS             complement(2149584..2150144)
                     /note="peptidyl-tRNA hydrolase"
                     /gene="BMETHMGA301857c"
     CDS             complement(2150264..2150899)
                     /note="50S ribosomal protein L25"
                     /gene="BMETHMGA301858c"
     CDS             complement(2151065..2152021)
                     /note="ribose-phosphate pyrophosphokinase"
                     /gene="BMETHMGA301859c"
     CDS             complement(2152042..2153415)
                     /note="UDP-N-acetylglucosamine pyrophosphorylase"
                     /gene="BMETHMGA301860c"
     CDS             complement(2153836..2154126)
                     /note="SpoVG family protein"
                     /gene="BMETHMGA301861c"
     CDS             complement(2154301..2154741)
                     /note="translation initiation inhibitor"
                     /gene="BMETHMGA301862c"
     CDS             complement(2154728..2155561)
                     /note="purine operon repressor"
                     /gene="BMETHMGA301863c"
     CDS             complement(2155621..2156403)
                     /note="4-diphosphocytidyl-2-C-methyl-D-erythritol
            kinase"
                     /gene="BMETHMGA301864c"
     CDS             complement(2157261..2157563)
                     /note="hypothetical protein"
                     /gene="BMETHMGA301866c"
     CDS             complement(2157733..2158437)
                     /note="sporulation peptidase YabG"
                     /gene="BMETHMGA301867c"
     CDS             complement(2158817..2159695)
                     /note="dimethyladenosine transferase"
                     /gene="BMETHMGA301868c"
     CDS             complement(2159688..2160248)
                     /note="primase/topoisomerase like protein"
                     /gene="BMETHMGA301869c"
     CDS             complement(2160509..2161690)
                     /note="hypothetical protein"
                     /gene="BMETHMGA301870c"
     CDS             complement(2161962..2162729)
                     /note="metal-dependent DNase"
                     /gene="BMETHMGA301871c"
```

```
CDS             complement(2162871..2164721)
                /note="methionyl-tRNA synthetase"
                /gene="BMETHMGA301872c"
CDS             2165469..2165759
                /note="transition state regulatory protein AbrB"
                /gene="BMETHMGA301873"
CDS             complement(2165861..2166736)
                /note="Uroporphyrin-III C/tetrapyrrole
       (Corrin/Porphyrin)
                methyltransferase"
                /gene="BMETHMGA301874c"
CDS             complement(2166708..2167001)
                /note="hypothetical protein"
                /gene="BMETHMGA301875c"
CDS             complement(2166988..2167731)
                /note="methyltransferase"
                /gene="BMETHMGA301876c"
CDS             complement(2167808..2168149)
                /note="DNA replication complex subunit"
                /gene="BMETHMGA301877c"
CDS             complement(2168322..2169149)
                /note="signal peptidase-like protein"
                /gene="BMETHMGA301878c"
CDS             complement(2169142..2170119)
                /note="DNA polymerase III, delta prime subunit"
                /gene="BMETHMGA301879c"
CDS             complement(2170163..2170603)
                /note="hypothetical protein"
                /gene="BMETHMGA301880c"
CDS             complement(2170627..2170956)
                /note="hypothetical protein yaaQ"
                /gene="BMETHMGA301881c"
CDS             complement(2171072..2171710)
                /note="thymidylate kinase"
                /gene="BMETHMGA301882c"
CDS             complement(2171761..2173197)
                /note="Orn/Lys/Arg decarboxylase major region"
                /gene="BMETHMGA301883c"
fasta_record    2173771..2186338
                /note="ConsensusfromContig43 Average coverage:
       536.63"
                /label=ConsensusfromContig43 Average coverage: 536.63
                /colour=10
CDS             complement(2173837..2175999)
                /note="ATP-dependent Clp protease-like protein"
                /gene="BMETHMGA301884c"
CDS             2176410..2177327
                /note="putative transporter"
                /gene="BMETHMGA301885"
CDS             complement(2177347..2178030)
                /note="hypothetical protein"
                /gene="BMETHMGA301886c"
CDS             complement(2178231..2178896)
```

```
                /note="hypothetical protein"
                /gene="BMETHMGA301887c"
CDS             2180577..2182295
                /note="phosphoenolpyruvate-protein
       phosphotransferase"
                /gene="BMETHMGA301888"
CDS             2182500..2183285
                /note="3-oxoacyl-(acyl-carrier protein) reductase"
                /gene="BMETHMGA301889"
CDS             2183408..2183851
                /note="hypothetical protein"
                /gene="BMETHMGA301890"
CDS             2184058..2184900
                /note="6-phosphogluconate dehydrogenase, NAD-binding
                subunit"
                /gene="BMETHMGA301891"
CDS             complement(2185027..2186184)
                /note="aminotransferase class I and II"
                /gene="BMETHMGA301892c"
CDS             2186285..2186797
                /note="ABC transporter, integral membrane subunit"
                /gene="BMETHMGA301893"
fasta_record    2186339..2188804
                /note="ConsensusfromContig44 Average coverage:
       731.03"
                /label=ConsensusfromContig44 Average coverage: 731.03
                /colour=11
CDS             2186821..2187822
                /note="NMT1/THI5 like domain-containing protein"
                /gene="BMETHMGA301894"
CDS             2187861..2188655
                /note="hydroxyethylthiazole kinase"
                /gene="BMETHMGA301895"
fasta_record    2188805..2196923
                /note="ConsensusfromContig45 Average coverage:
       506.96"
                /label=ConsensusfromContig45 Average coverage: 506.96
                /colour=10
CDS             2189016..2189615
                /note="precorrin-2 dehydrogenase"
                /gene="BMETHMGA301896"
CDS             2190201..2191397
                /note="8-amino-7-oxononanoate synthase"
                /gene="BMETHMGA301897"
CDS             2191394..2192137
                /note="biotin biosynthesis protein BioH"
                /gene="BMETHMGA301898"
CDS             2191948..2192979
                /note="Biotin biosynthesis protein BioC"
                /gene="BMETHMGA301899"
CDS             2193226..2193807
                /note="hypothetical protein"
                /gene="BMETHMGA301900"
```

```
CDS             2195397..2195936
                /note="spore germination protein, GerA"
                /gene="BMETHMGA301902"
CDS             2196060..2196404
                /note="spore germination protein"
                /gene="BMETHMGA301903"
CDS             2196504..2196803
                /note="spore germination protein"
                /gene="BMETHMGA301904"
fasta_record    2196924..2202767
                /note="ConsensusfromContig46 Average coverage:
       634.59"
                /label=ConsensusfromContig46 Average coverage: 634.59
                /colour=11
CDS             complement(2197264..2197593)
                /note="aminoglycoside phosphotransferase"
                /gene="BMETHMGA301906c"
CDS             2197758..2198231
                /note="hypothetical protein"
                /gene="BMETHMGA301907"
CDS             2198503..2198850
                /note="transcriptional regulator, MecI"
                /gene="BMETHMGA301908"
CDS             2198847..2199728
                /note="peptidase M56 BlaR1"
                /gene="BMETHMGA301909"
CDS             2199747..2200598
                /note="hypothetical protein"
                /gene="BMETHMGA301910"
CDS             complement(2201128..2201334)
                /note="DNA repair protein RadC"
                /gene="BMETHMGA301912c"
fasta_record    2202768..2206598
                /note="ConsensusfromContig47 Average coverage:
       629.51"
                /label=ConsensusfromContig47 Average coverage: 629.51
                /colour=10
CDS             complement(2203722..2204171)
                /note="hypothetical protein"
                /gene="BMETHMGA301913c"
CDS             2204645..2205004
                /note="nitrogen regulatory protein P-II"
                /gene="BMETHMGA301914"
CDS             2205010..2206314
                /note="ammonium transporter"
                /gene="BMETHMGA301915"
fasta_record    2206599..2246299
                /note="ConsensusfromContig48 Average coverage:
       493.67"
                /label=ConsensusfromContig48 Average coverage: 493.67
                /colour=11
CDS             complement(2206852..2207142)
                /note="Parvovirus coat protein VP-1 like protein"
```

```
                /gene="BMETHMGA301916c"
CDS             complement(2207784..2209142)
                /note="hypothetical protein"
                /gene="BMETHMGA301918c"
CDS             2209009..2210898
                /note="squalene/oxidosqualene cyclase"
                /gene="BMETHMGA301919"
CDS             complement(2210990..2213275)
                /note="hypothetical protein"
                /gene="BMETHMGA301920c"
CDS             2213637..2214287
                /note="hypothetical protein"
                /gene="BMETHMGA301921"
CDS             complement(2214643..2215185)
                /note="FMN reductase"
                /gene="BMETHMGA301922c"
CDS             complement(2216043..2216990)
                /note="Nitrialse"
                /gene="BMETHMGA301923c"
CDS             2217278..2218045
                /note="glycine/serine hydroxymethyltransferase"
                /gene="BMETHMGA301924"
CDS             complement(2218246..2219679)
                /note="nitrilotriacetate monooxygenase component
                A/pristinamycin IIA synthase subunit A"
                /gene="BMETHMGA301925c"
CDS             complement(2220238..2222406)
                /note="methyl-accepting chomotaxis sensory
       transducer"
                /gene="BMETHMGA301926c"
CDS             complement(2222836..2223612)
                /note="7-cyano-7-deazaguanosine(preQO) biosynthesis
                protein QueE"
                /gene="BMETHMGA301928c"
CDS             complement(2223557..2224039)
                /note="6-pyruvoyl tetrahydrobiopterin synthase"
                /gene="BMETHMGA301929c"
CDS             complement(2224045..2224704)
                /note="queuosine biosynthesis protein"
                /gene="BMETHMGA301930c"
CDS             2224997..2227027
                /note="formate dehydrogenase"
                /gene="BMETHMGA301931"
CDS             complement(2227172..2228521)
                /note="metal ion (Mn2+/Fe2+) transporter family metal
       ion
                transporter"
                /gene="BMETHMGA301932c"
CDS             complement(2228828..2230360)
                /note="bilirubin oxidase"
                /gene="BMETHMGA301933c"
CDS             2230620..2231840
                /note="tryptophan synthase, beta subunit"
```

```
                /gene="BMETHMGA301934"
CDS             complement(2232008..2233177)
                /note="NLP/P60 protein"
                /gene="BMETHMGA301935c"
CDS             complement(2233896..2234747)
                /note="hypothetical protein"
                /gene="BMETHMGA301936c"
CDS             complement(2235216..2235773)
                /note="hypothetical protein"
                /gene="BMETHMGA301938c"
CDS             complement(2235919..2237769)
                /note="hypothetical protein"
                /gene="BMETHMGA301939c"
CDS             complement(2237865..2238878)
                /note="type III restriction protein res subunit"
                /gene="BMETHMGA301940c"
CDS             complement(2239438..2240340)
                /note="hypothetical protein"
                /gene="BMETHMGA301941c"
CDS             2241189..2243015
                /note="serine protein kinase"
                /gene="BMETHMGA301942"
CDS             2243043..2244188
                /note="sporulation protein YhbH"
                /gene="BMETHMGA301943"
CDS             2244185..2245423
                /note="stage V sporulation protein R"
                /gene="BMETHMGA301944"
CDS             complement(2246150..2247049)
                /note="IS4 family transposase"
                /gene="BMETHMGA301946c"
fasta_record    2246300..2247188
                /note="ConsensusfromContig49 Average coverage:
       1,983.73"
                /label=ConsensusfromContig49 Average coverage:
       1,983.73
                /colour=10
fasta_record    2247189..2271422
                /note="ConsensusfromContig50 Average coverage:
       587.01"
                /label=ConsensusfromContig50 Average coverage: 587.01
                /colour=11
CDS             2247268..2247588
                /note="spore germination protein"
                /gene="BMETHMGA301947"
CDS             complement(2247662..2248834)
                /note="rod shape-determining protein RodA"
                /gene="BMETHMGA301948c"
CDS             2249087..2249947
                /note="2,5-didehydrogluconate reductase"
                /gene="BMETHMGA301949"
CDS             2250433..2251476
                /note="hypothetical protein"
```

```
                     /gene="BMETHMGA301950"
     CDS             2251473..2252060
                     /note="hypothetical protein"
                     /gene="BMETHMGA301951"
     CDS             2252327..2252971
                     /note="NAD-dependent epimerase/dehydratase"
                     /gene="BMETHMGA301952"
     CDS             complement(2253031..2253987)
                     /note="cadmium, cobalt and zinc/H(+)-K(+) antiporter"
                     /gene="BMETHMGA301953c"
     CDS             complement(2254034..2254414)
                     /note="cadmium efflux system accessory protein"
                     /gene="BMETHMGA301954c"
     CDS             2255380..2256330
                     /note="divalent cation transport protein"
                     /gene="BMETHMGA301956"
     CDS             complement(2256436..2257230)
                     /note="short-chain dehydrogenase/reductase family
              protein"
                     /gene="BMETHMGA301957c"
     CDS             2257444..2257809
                     /note="hypothetical protein"
                     /gene="BMETHMGA301958"
     CDS             2257763..2259649
                     /note="ATP-binding region ATPase domain protein"
                     /gene="BMETHMGA301959"
     CDS             2259869..2260618
                     /note="transcriptional activator TipA"
                     /gene="BMETHMGA301960"
     CDS             2260762..2261613
                     /note="glyoxalase family protein"
                     /gene="BMETHMGA301961"
     CDS             complement(2261654..2262595)
                     /note="manganase-dependent inorganin pyrophosphatase"
                     /gene="BMETHMGA301962c"
     CDS             complement(2262815..2263636)
                     /note="NH(3)-dependent NAD synthetase"
                     /gene="BMETHMGA301964c"
     CDS             2263868..2264278
                     /note="BadM/Rrf2 family transcriptional regulator"
                     /gene="BMETHMGA301965"
     CDS             2264332..2265444
                     /note="tRNA
                     (5-methylaminomethyl-2-thiouridylate)-
              methyltransferase"
                     /gene="BMETHMGA301966"
     CDS             2265578..2266528
                     /note="luciferase-type oxidoreductase"
                     /gene="BMETHMGA301967"
     CDS             complement(2266738..2267304)
                     /note="hypothetical protein"
                     /gene="BMETHMGA301968c"
     CDS             complement(2267571..2267861)
```

```
                /note="hypothetical protein ydzA"
                /gene="BMETHMGA301969c"
CDS             complement(2268104..2269006)
                /note="transcriptional regulator, LysR family"
                /gene="BMETHMGA301970c"
CDS             2269135..2271135
                /note="phage infection protein"
                /gene="BMETHMGA301971"
fasta_record    2271423..2273472
                /note="ConsensusfromContig51 Average coverage:
       1,401.02"
                /label=ConsensusfromContig51 Average coverage:
       1,401.02
                /colour=10
CDS             2271512..2271919
                /note="spore germination protein"
                /gene="BMETHMGA301972"
fasta_record    2273473..2278792
                /note="ConsensusfromContig52 Average coverage:
       516.50"
                /label=ConsensusfromContig52 Average coverage: 516.50
                /colour=11
CDS             complement(2273562..2274044)
                /note="hypothetical protein"
                /gene="BMETHMGA301974c"
CDS             2274227..2275000
                /note="3-ketoacyl-(acyl-carrier-protein) reductase"
                /gene="BMETHMGA301975"
CDS             complement(2275078..2276487)
                /note="multidrug efflux protein"
                /gene="BMETHMGA301976c"
CDS             complement(2276631..2277227)
                /note="ribosomal-protein-alanine acetyltransferase"
                /gene="BMETHMGA301977c"
CDS             complement(2277497..2278624)
                /note="imidazolonepropionase"
                /gene="BMETHMGA301978c"
fasta_record    2278793..2311810
                /note="ConsensusfromContig53 Average coverage:
       609.85"
                /label=ConsensusfromContig53 Average coverage: 609.85
                /colour=10
CDS             2279396..2279851
                /note="hypothetical protein"
                /gene="BMETHMGA301980"
CDS             2279848..2280546
                /note="peptidase M22 glycoprotease"
                /gene="BMETHMGA301981"
CDS             2280572..2281021
                /note="ribosomal-protein-alanine acetyltransferase"
                /gene="BMETHMGA301982"
CDS             2281018..2282040
                /note="metalloendopeptidase, glycoprotease family"
```

```
                /gene="BMETHMGA301983"
CDS             complement(2282294..2284213)
                /note="ABC transporter ATP-binding protein"
                /gene="BMETHMGA301984c"
CDS             2284354..2284848
                /note="molybdenum cofactor biosynthesis protein C"
                /gene="BMETHMGA301985"
CDS             2284870..2285517
                /note="redox-sensing transcriptional repressor Rex"
                /gene="BMETHMGA301986"
CDS             complement(2285840..2286529)
                /note="putative membranse protease"
                /gene="BMETHMGA301987c"
CDS             2286846..2287148
                /note="10 kDa chaperonin"
                /gene="BMETHMGA301988"
CDS             2287241..2288854
                /note="Chaperonin GroEL"
                /gene="BMETHMGA301989"
CDS             2289272..2289985
                /note="hypothetical protein"
                /gene="BMETHMGA301990"
CDS             2290066..2291217
                /note="Flp pilus assembly ATpase CpaE-like"
                /gene="BMETHMGA301991"
CDS             2291230..2292582
                /note="Flp pilus assembly protein ATPase CpaF"
                /gene="BMETHMGA301992"
CDS             2292579..2293505
                /note="hypothetical protein"
                /gene="BMETHMGA301993"
CDS             2293510..2294442
                /note="type II secretion system protein"
                /gene="BMETHMGA301994"
CDS             2295077..2296099
                /note="D-amino acid dehydrogenase, large subunit"
                /gene="BMETHMGA301995"
CDS             2297796..2298356
                /note="hypothetical protein"
                /gene="BMETHMGA301997"
CDS             2300222..2300938
                /note="hypothetical protein"
                /gene="BMETHMGA302002"
CDS             2300947..2301729
                /note="D-amino-acid dehydrogenase"
                /gene="BMETHMGA302003"
CDS             2302379..2303773
                /note="putative integral inner membrane protein"
                /gene="BMETHMGA302004"
CDS             2303795..2304832
                /note="ribonuclease III"
                /gene="BMETHMGA302005"
CDS             2304786..2305211
```

```
                /note="hypothetical protein"
                /gene="BMETHMGA302006"
CDS             2305339..2306415
                /note="methanol dehydrogenase regulatory protein"
                /gene="BMETHMGA302007"
CDS             2306418..2307635
                /note="hypothetical protein"
                /gene="BMETHMGA302008"
CDS             2307632..2309842
                /note="hypothetical protein"
                /gene="BMETHMGA302009"
CDS             2310185..2311705
                /note="bifuntional GMP synthase/glutamine
       amidotransferase
                protein"
                /gene="BMETHMGA302010"
fasta_record    2311811..2327295
                /note="ConsensusfromContig54 Average coverage:
       662.83"
                /label=ConsensusfromContig54 Average coverage: 662.83
                /colour=11
CDS             complement(2312045..2312803)
                /note="hypothetical protein"
                /gene="BMETHMGA302011c"
CDS             2313115..2314581
                /note="inosine 5'-monophosphate dehydrogenase"
                /gene="BMETHMGA302012"
CDS             2314732..2316129
                /note="serine-type D-Ala-D-Ala carboxypeptidase"
                /gene="BMETHMGA302015"
CDS             2316279..2317160
                /note="pyridoxine biosynthesis protein"
                /gene="BMETHMGA302018"
CDS             2317172..2317762
                /note="pyridoxine biosynthesis amidotransferase"
                /gene="BMETHMGA302019"
CDS             2318076..2319395
                /note="seryl-tRNA synthetase"
                /gene="BMETHMGA302020"
CDS             complement(2319782..2320450)
                /note="deoxynyucleoside kinase"
                /gene="BMETHMGA302022c"
CDS             complement(2320428..2320913)
                /note="deoxyguanosine kinase"
                /gene="BMETHMGA302023c"
CDS             complement(2321297..2322622)
                /note="spore peptidoglycan hydrolase"
                /gene="BMETHMGA302024c"
CDS             2322733..2323269
                /note="zinc binding CMP/dCMP deaminase"
                /gene="BMETHMGA302025"
CDS             2324161..2325468
                /note="DNA polymerase III subunits gamma and tau"
```

```
                /gene="BMETHMGA302026"
CDS             2325485..2325808
                /note="hypothetical protein"
                /gene="BMETHMGA302027"
CDS             2325824..2326420
                /note="recombination protein RecR"
                /gene="BMETHMGA302028"
fasta_record    2327296..2371709
                /note="ConsensusfromContig55 Average coverage:
       523.11"
                /label=ConsensusfromContig55 Average coverage: 523.11
                /colour=10
CDS             2327446..2327916
                /note="hypothetical protein"
                /gene="BMETHMGA302029"
CDS             2327974..2328564
                /note="Rieske(2Fe-2S) iron-sulfur domain protein"
                /gene="BMETHMGA302030"
CDS             2328745..2329242
                /note="cytochrome b6"
                /gene="BMETHMGA302031"
CDS             2329304..2330071
                /note="menaquinol-cytochrome c reductase"
                /gene="BMETHMGA302032"
CDS             2330149..2330748
                /note="hypothetical protein"
                /gene="BMETHMGA302033"
CDS             2330825..2331619
                /note="sporulation protein YpjB"
                /gene="BMETHMGA302034"
CDS             2331746..2332420
                /note="peptidase membrane zinc metallopeptidase"
                /gene="BMETHMGA302035"
CDS             complement(2332448..2333236)
                /note="hypothetical protein"
                /gene="BMETHMGA302036c"
CDS             2333530..2333862
                /note="hypothetical protein"
                /gene="BMETHMGA302037"
CDS             2333876..2334676
                /note="dihyrodipicolinate reductase"
                /gene="BMETHMGA302038"
CDS             2334689..2335099
                /note="methylglyoxal synthase"
                /gene="BMETHMGA302039"
CDS             2335104..2335820
                /note="LmbE family protein"
                /gene="BMETHMGA302040"
CDS             2335817..2336971
                /note="N-acetyl-alpha-D-glucosaminyl L-malate
       synthase"
                /gene="BMETHMGA302041"
CDS             2336952..2338151
```

```
                /note="tRNa nucleotidyltransferase"
                /gene="BMETHMGA302042"
CDS             2338124..2339125
                /note="biotin/acetyl-CoA-carboyxlase ligase"
                /gene="BMETHMGA302043"
CDS             2339352..2340188
                /note="3-methyl-2-oxobutanoate
       hydroxymethyltransferase"
                /gene="BMETHMGA302044"
CDS             2340185..2341036
                /note="pantoate/beta-alanine ligase"
                /gene="BMETHMGA302045"
CDS             2341049..2341432
                /note="aspartate alpha-decarboxylase"
                /gene="BMETHMGA302046"
CDS             2341557..2344367
                /note="DNAQ family exonuclease/DinG family helicase"
                /gene="BMETHMGA302047"
CDS             2344722..2345264
                /note="propeptide PepSY and peptidase M4"
                /gene="BMETHMGA302048"
CDS             2345284..2346483
                /note="aspartate aminotransferase"
                /gene="BMETHMGA302049"
CDS             2346984..2347937
                /note="asparaginyl-tRNA synthetase"
                /gene="BMETHMGA302050"
CDS             2347997..2348806
                /note="chromosome replication initiation protein"
                /gene="BMETHMGA302051"
CDS             2348870..2349526
                /note="endonuclease III"
                /gene="BMETHMGA302052"
CDS             2349532..2350038
                /note="hypothetical protein"
                /gene="BMETHMGA302053"
CDS             complement(2350072..2353047)
                /note="penicillin-binding protein, 1A family"
                /gene="BMETHMGA302054c"
CDS             complement(2353028..2353639)
                /note="Holliday junction-specific endonuclease"
                /gene="BMETHMGA302055c"
CDS             2353758..2354744
                /note="hypothetical protein"
                /gene="BMETHMGA302056"
CDS             2355260..2355628
                /note="hypothetical protein"
                /gene="BMETHMGA302057"
CDS             2356185..2356724
                /note="hyppothetical protein"
                /gene="BMETHMGA302058"
CDS             complement(2356884..2357357)
                /note="heat shock protein class I"
```

```
                 /gene="BMETHMGA302059c"
CDS              2357512..2358840
                 /note="hybrid cluster protein/hydroxylamine
       reductase"
                 /gene="BMETHMGA302060"
CDS              complement(2358902..2359594)
                 /note="cyclic nucleotide-binding protein"
                 /gene="BMETHMGA302061c"
CDS              2359827..2362094
                 /note="putative ATP-dependent helicase"
                 /gene="BMETHMGA302062"
CDS              2362081..2363373
                 /note="putative nucleic acid binding protein"
                 /gene="BMETHMGA302063"
CDS              2363711..2364163
                 /note="spore coat protein D"
                 /gene="BMETHMGA302064"
CDS              2364451..2364753
                 /note="DivIVA domain protein"
                 /gene="BMETHMGA302065"
CDS              2365356..2367416
                 /note="PAS/PAC sensor-containing diguanylate
                 cyclase/phosphodiesterase"
                 /gene="BMETHMGA302066"
CDS              2367429..2368568
                 /note="rRNA (quanine-N(2)-)-methyltransferase"
                 /gene="BMETHMGA302067"
CDS              2369613..2371547
                 /note="ATP-dependent helicase"
                 /gene="BMETHMGA302068"
fasta_record     2371710..2494177
                 /note="ConsensusfromContig56 Average coverage:
       636.97"
                 /label=ConsensusfromContig56 Average coverage: 636.97
                 /colour=11
CDS              2371830..2372972
                 /note="ATPase"
                 /gene="BMETHMGA302069"
CDS              2372999..2373649
                 /note="hypothetical protein"
                 /gene="BMETHMGA302070"
CDS              2373749..2374480
                 /note="DNA alkylation repair enzyme"
                 /gene="BMETHMGA302071"
CDS              2374590..2375678
                 /note="glucose dehydrogenase-B"
                 /gene="BMETHMGA302072"
CDS              complement(2375814..2376467)
                 /note="transcriptional regulator. TetR family"
                 /gene="BMETHMGA302073c"
CDS              complement(2376534..2377562)
                 /note="ABC transporter, permease"
                 /gene="BMETHMGA302074c"
```

```
CDS             complement(2377559..2378290)
                /note="ABC transproter, ATP-binding protein"
                /gene="BMETHMGA302075c"
CDS             2378592..2379122
                /note="hemerythrin HHE cation binding domain-
       containig
                protein"
                /gene="BMETHMGA302076"
CDS             complement(2379242..2379808)
                /note="LmbE family protein"
                /gene="BMETHMGA302077c"
CDS             complement(2379967..2380314)
                /note="hypothetical protein"
                /gene="BMETHMGA302078c"
CDS             complement(2380729..2381562)
                /note="phosphomethylpyrimidine kinase"
                /gene="BMETHMGA302079c"
CDS             2381897..2382448
                /note="hypothetical protein"
                /gene="BMETHMGA302080"
CDS             2382625..2383089
                /note="molybdopterin biosynthesis MoaE protein"
                /gene="BMETHMGA302081"
CDS             2383339..2384034
                /note="uracil-DNa glycosylase"
                /gene="BMETHMGA302082"
CDS             2384775..2385146
                /note="hypothetical protein"
                /gene="BMETHMGA302083"
CDS             complement(2385249..2385734)
                /note="spore coat protein GerQ"
                /gene="BMETHMGA302084c"
CDS             complement(2386122..2386868)
                /note="chlorite dismutase"
                /gene="BMETHMGA302085c"
CDS             2387165..2388136
                /note="phosphate acetyltransferase"
                /gene="BMETHMGA302086"
CDS             2388277..2389170
                /note="lipoate-protein ligase A"
                /gene="BMETHMGA302087"
CDS             complement(2389372..2390076)
                /note="transcriptional factor, RsfA family"
                /gene="BMETHMGA302088c"
CDS             2390806..2392173
                /note="HD superfamily phosphohydrolase"
                /gene="BMETHMGA302089"
CDS             2392272..2392757
                /note="hypothetical protein"
                /gene="BMETHMGA302090"
CDS             2393308..2393811
                /note="metal-dependent protease"
                /gene="BMETHMGA302091"
```

```
CDS             2393849..2394361
                /note="hypothetical protein"
                /gene="BMETHMGA302092"
CDS             complement(2394528..2396594)
                /note="penicillin binding protein, 1A family"
                /gene="BMETHMGA302093c"
CDS             2396825..2397652
                /note="spermidine synthase"
                /gene="BMETHMGA302094"
CDS             2397738..2398610
                /note="agmatinase"
                /gene="BMETHMGA302095"
CDS             2398877..2399290
                /note="hypothetical protein"
                /gene="BMETHMGA302096"
CDS             2399287..2400957
                /note="arginyl-tRNA synthetase"
                /gene="BMETHMGA302097"
CDS             complement(2401323..2402528)
                /note="cardiolipin synthetase domain-protein"
                /gene="BMETHMGA302098c"
CDS             2402725..2404860
                /note="ferredoxin, 4Fe-4S"
                /gene="BMETHMGA302099"
CDS             2405144..2406331
                /note="acetyl-CoA acetyltransferase"
                /gene="BMETHMGA302100"
CDS             2406430..2407263
                /note="3-hydroxybutyryl-CoA dehydrogenase"
                /gene="BMETHMGA302101"
CDS             2407340..2408470
                /note="acyl-CoA dehydrogenase"
                /gene="BMETHMGA302102"
CDS             2408495..2409634
                /note="acyl-CoA dehydrogenase"
                /gene="BMETHMGA302103"
CDS             2409685..2410332
                /note="transcriptional regulator, TetR family"
                /gene="BMETHMGA302104"
CDS             2410325..2413576
                /note="methylmalonyl-CoA mutase, large subunit"
                /gene="BMETHMGA302105"
CDS             2414348..2414887
                /note="DNA-directed RNA polymerase, delta subunit"
                /gene="BMETHMGA302107"
CDS             2415113..2416717
                /note="CTP synthase"
                /gene="BMETHMGA302108"
CDS             complement(2416793..2417329)
                /note="hypothetical protein"
                /gene="BMETHMGA302109c"
CDS             2417536..2417907
                /note="two-component response regulator"
```

```
                /gene="BMETHMGA302110"
CDS             2418148..2419005
                /note="fructose-biphosphate aldolase"
                /gene="BMETHMGA302111"
CDS             2419998..2421284
                /note="UDP-N-acetylglucoseamine 1-
       carboxyvinyltransferase"
                /gene="BMETHMGA302112"
CDS             2421339..2422304
                /note="fructose-1,6-biphosphatase, class II"
                /gene="BMETHMGA302113"
CDS             2422758..2424026
                /note="transcriptional termination factor Rho"
                /gene="BMETHMGA302114"
CDS             2424677..2425255
                /note="thymidine kinase"
                /gene="BMETHMGA302115"
CDS             2425378..2426457
                /note="peptide chain release factor 1"
                /gene="BMETHMGA302116"
CDS             2426417..2427319
                /note="protein methyltransferase, release factor-
       specific"
                /gene="BMETHMGA302117"
CDS             2427457..2428137
                /note="hypothetical protein"
                /gene="BMETHMGA302118"
CDS             2428377..2429423
                /note="Sua5/YciO/YrdC/YwlC family protein"
                /gene="BMETHMGA302119"
CDS             2429534..2430088
                /note="hypothetical protein"
                /gene="BMETHMGA302120"
CDS             2430233..2430718
                /note="protein-tyrosine phosphatase"
                /gene="BMETHMGA302121"
CDS             2430693..2431991
                /note="methyl-accepting chemotaxis protein"
                /gene="BMETHMGA302122"
CDS             2432149..2432595
                /note="ribose-5-phosphate isomerase B"
                /gene="BMETHMGA302123"
CDS             2432791..2433360
                /note="hypothetical protein"
                /gene="BMETHMGA302124"
CDS             2433488..2434759
                /note="serine hydroxymethyltransferase"
                /gene="BMETHMGA302125"
CDS             2434979..2435608
                /note="uracil phosphoribosyltransferase"
                /gene="BMETHMGA302126"
CDS             2436056..2437027
                /note="NERD domain protein"
```

```
                     /gene="BMETHMGA302127"
CDS                  2437338..2439530
                     /note="minor extracellular serine protease"
                     /gene="BMETHMGA302128"
CDS                  2440217..2440591
                     /note="ATP synthase F0 subunit I"
                     /gene="BMETHMGA302129"
CDS                  2440614..2441324
                     /note="ATP synthase F0F1 subunit A"
                     /gene="BMETHMGA302130"
CDS                  2441856..2442302
                     /note="ATP synthase subunit B"
                     /gene="BMETHMGA302131"
CDS                  2442299..2442835
                     /note="ATP synthase F1, subunit delta"
                     /gene="BMETHMGA302132"
CDS                  2442860..2444368
                     /note="ATP synthase F0F1 subunit alpha"
                     /gene="BMETHMGA302133"
CDS                  2444477..2445265
                     /note="ATP synthase F0F1 gamma subunit"
                     /gene="BMETHMGA302134"
CDS                  2445311..2446795
                     /note="ATP synthase F0F1 subunit beta"
                     /gene="BMETHMGA302135"
CDS                  2446818..2447225
                     /note="ATP synthase F0F1 subunit epsilon"
                     /gene="BMETHMGA302136"
CDS                  2447607..2448041
                     /note="NADH dehydrogenase, subunit A"
                     /gene="BMETHMGA302137"
CDS                  2448011..2448544
                     /note="NADH dehydrogenase subunit B"
                     /gene="BMETHMGA302138"
CDS                  2448548..2449513
                     /note="NADH dehydrogenase subunit c"
                     /gene="BMETHMGA302139"
CDS                  2449535..2450617
                     /note="NADH dehydrogenase subunit D"
                     /gene="BMETHMGA302140"
CDS                  2450617..2451624
                     /note="NADH dehydrogenase subunit H"
                     /gene="BMETHMGA302141"
CDS                  2451665..2452084
                     /note="NADH dehydrogenase subunit I"
                     /gene="BMETHMGA302142"
CDS                  2452081..2452602
                     /note="NADH dehydrogenase subunit J"
                     /gene="BMETHMGA302143"
CDS                  2452599..2452913
                     /note="NADH dehydrogenase subunit K"
                     /gene="BMETHMGA302144"
CDS                  2453005..2454918
```

```
                /note="NADH dehydrogenase subunit L"
                /gene="BMETHMGA302145"
CDS             2454918..2456432
                /note="NADH dehydrogenase subunit M"
                /gene="BMETHMGA302146"
CDS             2456439..2457944
                /note="NADH dehydrogenase subunit N"
                /gene="BMETHMGA302147"
CDS             2458190..2458462
                /note="hypothetical protein"
                /gene="BMETHMGA302148"
CDS             2458685..2459419
                /note="hypothetical protein"
                /gene="BMETHMGA302149"
CDS             2459725..2460762
                /note="UDP-N-acetylglucosamine 1-
       carboxyvinyltransferase"
                /gene="BMETHMGA302150"
CDS             2461013..2462059
                /note="stage II sporulation protein D"
                /gene="BMETHMGA302151"
CDS             2462588..2463439
                /note="ATP-binding region ATPase domain protein"
                /gene="BMETHMGA302152"
CDS             complement(2464380..2464832)
                /note="hypothetical protein"
                /gene="BMETHMGA302153c"
CDS             2464953..2465957
                /note="SpoIIQ"
                /gene="BMETHMGA302154"
CDS             2466288..2466668
                /note="sporulation stage III transcriptional
       regulator
                SpoIIID"
                /gene="BMETHMGA302155"
CDS             2467194..2468195
                /note="cell shape determining protein, MreB/Mr1
       family"
                /gene="BMETHMGA302156"
CDS             2468366..2469205
                /note="flagellar basal-body rod protein"
                /gene="BMETHMGA302157"
CDS             2469256..2470098
                /note="flagellar hook-basal body protein"
                /gene="BMETHMGA302158"
CDS             2470549..2470983
                /note="(3R)-hydroxmyristoyl dehydratase"
                /gene="BMETHMGA302159"
CDS             complement(2471862..2472305)
                /note="hypothetical protein"
                /gene="BMETHMGA302161c"
CDS             2473414..2474097
                /note="putative membrane protein"
```

```
                     /gene="BMETHMGA302163"
CDS                  2474501..2475946
                     /note="Aldehyde/NADP-dependent glyceraldehyde-3-
         phosphate
                     dehydrogenase"
                     /gene="BMETHMGA302164"
CDS                  2476529..2477020
                     /note="hypothetical protein"
                     /gene="BMETHMGA302165"
CDS                  2477111..2477539
                     /note="hypothetical protein"
                     /gene="BMETHMGA302166"
CDS                  complement(2477734..2480544)
                     /note="hypothetical protein"
                     /gene="BMETHMGA302167c"
CDS                  complement(2480537..2482159)
                     /note="putative replication initiation protein"
                     /gene="BMETHMGA302168c"
CDS                  complement(2482461..2482907)
                     /note="hypothetical protein"
                     /gene="BMETHMGA302169c"
CDS                  complement(2483375..2484247)
                     /note="hypothetical protein"
                     /gene="BMETHMGA302170c"
CDS                  2484734..2485351
                     /note="hypothetical protein"
                     /gene="BMETHMGA302171"
CDS                  complement(2485585..2486418)
                     /note="hypothetical protein"
                     /gene="BMETHMGA302172c"
CDS                  2486638..2488023
                     /note="carboxyl-terminal protease"
                     /gene="BMETHMGA302173"
CDS                  2488007..2490391
                     /note="hypothetical protein"
                     /gene="BMETHMGA302174"
CDS                  2491035..2493743
                     /note="surface layer glycoprotein SgsE precursor"
                     /gene="BMETHMGA302175"
fasta_record         2494178..2597108
                     /note="ConsensusfromContig57 Average coverage:
         618.23"
                     /label=ConsensusfromContig57 Average coverage: 618.23
                     /colour=10
CDS                  2494388..2494873
                     /note="hypothetical protein"
                     /gene="BMETHMGA302176"
CDS                  complement(2495141..2495887)
                     /note="ABC-type polar amino acid transport system,
         ATPase
                     component"
                     /gene="BMETHMGA302177c"
CDS                  complement(2495902..2496546)
```

```
                     /note="L-cystine ABC transporter, permease"
                     /gene="BMETHMGA302178c"
     CDS             complement(2496574..2497407)
                     /note="cystine-binding periplasmic protein precursor"
                     /gene="BMETHMGA302179c"
     CDS             complement(2497446..2498024)
                     /note="transcriptional regulator"
                     /gene="BMETHMGA302180c"
     CDS             complement(2498336..2499307)
                     /note="aldo/keto reductase"
                     /gene="BMETHMGA302181c"
     CDS             complement(2500508..2501740)
                     /note="major facilitator superfamily protein"
                     /gene="BMETHMGA302182c"
     CDS             complement(2501965..2502816)
                     /note="hypothetical protein"
                     /gene="BMETHMGA302183c"
     CDS             2503981..2504460
                     /note="hypothetical protein"
                     /gene="BMETHMGA302185"
     CDS             complement(2504852..2505700)
                     /note="linocin_M18 bacteriocin family protein"
                     /gene="BMETHMGA302186c"
     CDS             complement(2505725..2506219)
                     /note="hypothetical protein"
                     /gene="BMETHMGA302187c"
     CDS             2506658..2508217
                     /note="copper resistance protein"
                     /gene="BMETHMGA302188"
     CDS             2508245..2508910
                     /note="hypothetical protein"
                     /gene="BMETHMGA302189"
     CDS             complement(2509175..2510350)
                     /note="toxic anion resistance family protein"
                     /gene="BMETHMGA302190c"
     CDS             complement(2510347..2511042)
                     /note="putative phosphatase"
                     /gene="BMETHMGA302191c"
     CDS             complement(2511356..2511883)
                     /note="hypothetical protein"
                     /gene="BMETHMGA302192c"
     CDS             complement(2512316..2513314)
                     /note="VanW family protein"
                     /gene="BMETHMGA302193c"
     CDS             complement(2513698..2515239)
                     /note="multidrug resistance protein"
                     /gene="BMETHMGA302194c"
     CDS             complement(2515514..2516632)
                     /note="oxidoreductase yurR"
                     /gene="BMETHMGA302195c"
     CDS             complement(2517942..2518280)
                     /note="sortase family protein"
                     /gene="BMETHMGA302196c"
```

```
CDS             complement(2518641..2520182)
                /note="putative cell surface protein"
                /gene="BMETHMGA302197c"
CDS             complement(2520125..2521678)
                /note="hypothetical protein"
                /gene="BMETHMGA302200c"
CDS             complement(2522726..2524207)
                /note="glutamate synthase, subunit beta"
                /gene="BMETHMGA302201c"
CDS             complement(2524258..2528820)
                /note="glutamate synthase"
                /gene="BMETHMGA302202c"
CDS             2528945..2529847
                /note="transcriptional regulator, LysR family"
                /gene="BMETHMGA302203"
CDS             2530358..2530903
                /note="hypothetical protein"
                /gene="BMETHMGA302204"
CDS             complement(2531697..2533238)
                /note="malate:quinone oxidoreductase"
                /gene="BMETHMGA302205c"
CDS             complement(2533733..2534176)
                /note="hypothetical protein"
                /gene="BMETHMGA302206c"
CDS             2536551..2536856
                /note="recombinase"
                /gene="BMETHMGA302208"
CDS             2536901..2538082
                /note="site-specific recombinase"
                /gene="BMETHMGA302209"
CDS             complement(2538094..2539476)
                /note="RNA methyltransferase, TrmA family"
                /gene="BMETHMGA302210c"
CDS             complement(2540138..2540806)
                /note="methionine aminopeptidase"
                /gene="BMETHMGA302211c"
CDS             complement(2541223..2542974)
                /note="phosphodiesterase/alkaline phosphatase D"
                /gene="BMETHMGA302212c"
CDS             complement(2543125..2544951)
                /note="amino acid permease"
                /gene="BMETHMGA302213c"
CDS             complement(2545354..2546208)
                /note="NLPA lipoprotein"
                /gene="BMETHMGA302214c"
CDS             complement(2546276..2547490)
                /note="acyl-CoA dehydrogenase, type 2"
                /gene="BMETHMGA302215c"
CDS             complement(2547474..2549462)
                /note="acetoacetyl-CoA synthase"
                /gene="BMETHMGA302216c"
CDS             complement(2549465..2550631)
                /note="acetyl-CoA acetyltransferase"
```

```
                /gene="BMETHMGA302217c"
CDS             complement(2551658..2553067)
                /note="nicotinate phosphoribosyltransferase"
                /gene="BMETHMGA302218c"
CDS             complement(2553573..2554535)
                /note="intracellular serine protease"
                /gene="BMETHMGA302219c"
CDS             complement(2554763..2555686)
                /note="diacylglycerol kinase"
                /gene="BMETHMGA302220c"
CDS             complement(2555941..2556771)
                /note="undecaprenol kinase"
                /gene="BMETHMGA302221c"
CDS             complement(2557344..2558210)
                /note="fructosamine/Ketosamine-3-kinase"
                /gene="BMETHMGA302222c"
CDS             complement(2558207..2558677)
                /note="protein tyrosine phosphatase"
                /gene="BMETHMGA302223c"
CDS             2558978..2559589
                /note="hypothetical protein"
                /gene="BMETHMGA302224"
CDS             2559780..2560292
                /note="regulatory protein MerR"
                /gene="BMETHMGA302225"
CDS             complement(2560620..2562407)
                /note="thiamine biosynthesis protein ThiC"
                /gene="BMETHMGA302226c"
CDS             complement(2562484..2563053)
                /note="ThiW protein"
                /gene="BMETHMGA302227c"
CDS             complement(2563753..2565180)
                /note="aspartyl/glutamyl-tRNA amidotransferase
       subunit B"
                /gene="BMETHMGA302228c"
CDS             complement(2565193..2566431)
                /note="aspartyl/glutamyl-tRNA amidotransferase,
       subunit B"
                /gene="BMETHMGA302229c"
CDS             complement(2566663..2566953)
                /note="aspartyl/glutamyl-tRNA amidotransferase,
       subunit C"
                /gene="BMETHMGA302230c"
CDS             complement(2567246..2567722)
                /note="hypothetical protein"
                /gene="BMETHMGA302231c"
CDS             complement(2567885..2569168)
                /note="isocitrate lyase"
                /gene="BMETHMGA302232c"
CDS             complement(2569205..2570797)
                /note="malate synthase A"
                /gene="BMETHMGA302233c"
CDS             complement(2571055..2572602)
```

```
                 /note="1-pyrroline-5-carboxylate dehydrogenase"
                 /gene="BMETHMGA302234c"
CDS              complement(2572760..2573767)
                 /note="CamS sex pheromone cAM373 family protein"
                 /gene="BMETHMGA302235c"
CDS              complement(2573946..2575952)
                 /note="DNA ligse"
                 /gene="BMETHMGA302236c"
CDS              complement(2576008..2578212)
                 /note="ATP-dependent DNA helicase PcrA"
                 /gene="BMETHMGA302237c"
CDS              complement(2578287..2578973)
                 /note="geranylgeranylglyceryl phosphate synthase-like
                 protein"
                 /gene="BMETHMGA302238c"
CDS              complement(2579196..2579507)
                 /note="hypothetical protein"
                 /gene="BMETHMGA302239c"
CDS              complement(2579540..2580604)
                 /note="hypothetical protein"
                 /gene="BMETHMGA302240c"
CDS              complement(2580658..2582271)
                 /note="adenine deaminase"
                 /gene="BMETHMGA302241c"
CDS              complement(2583069..2584232)
                 /note="phosphoribosylamine-glycine ligase"
                 /gene="BMETHMGA302242c"
CDS              complement(2584454..2585992)
                 /note="bifucntional
                 phosphoribosylaminoimidazolecarboxamide
                 formyltransferase/IMP cyclohydrolase"
                 /gene="BMETHMGA302243c"
CDS              complement(2585971..2586573)
                 /note="phosphoribosylglycinamide formyltransferase"
                 /gene="BMETHMGA302244c"
CDS              complement(2586590..2587615)
                 /note="phosphoribosylformylglycinamidine cyclo-
       ligase"
                 /gene="BMETHMGA302245c"
CDS              complement(2587689..2589104)
                 /note="amidophosphoribosyltransferase"
                 /gene="BMETHMGA302246c"
CDS              complement(2589089..2591308)
                 /note="phosphoribosylformylglycinamidine synthase
       subunit
                 II"
                 /gene="BMETHMGA302247c"
CDS              complement(2591292..2591930)
                 /note="phosphoribosylformylglycinamidine synthase
       subunit
                 I"
                 /gene="BMETHMGA302248c"
CDS              complement(2592217..2592942)
```

```
                /note="phosphoribosylaminoimidazole-
         succinocarboxamide
                synthase"
                /gene="BMETHMGA302249c"
CDS             complement(2593021..2594313)
                /note="adenylosuccinate lyase"
                /gene="BMETHMGA302250c"
CDS             complement(2594317..2595468)
                /note="phosphoribosylaminoimidazole carboxylase"
                /gene="BMETHMGA302251c"
CDS             complement(2595455..2595943)
                /note="phosphoribosylaminoimidazole carboyxlase
         catalytic
                subunit PurE"
                /gene="BMETHMGA302252c"
fasta_record    2597109..2664328
                /note="ConsensusfromContig58 Average coverage:
         587.33"
                /label=ConsensusfromContig58 Average coverage: 587.33
                /colour=11
CDS             complement(2597277..2598290)
                /note="phosphate ABC transporter, permease protein
         PstA"
                /gene="BMETHMGA302253c"
CDS             complement(2598203..2599096)
                /note="phosphate ABC transporter, permease protein
         PstC"
                /gene="BMETHMGA302254c"
CDS             complement(2599173..2600126)
                /note="phosphate binding protein"
                /gene="BMETHMGA302255c"
CDS             2600623..2600964
                /note="cell wall hyrolase"
                /gene="BMETHMGA302256"
CDS             complement(2601099..2601776)
                /note="beta-phosphoglucomutase"
                /gene="BMETHMGA302257c"
CDS             complement(2601903..2602841)
                /note="glycoside hyrolase family protein"
                /gene="BMETHMGA302258c"
CDS             complement(2602949..2604214)
                /note="glycoside hydrolase family protein"
                /gene="BMETHMGA302259c"
CDS             complement(2604233..2605066)
                /note="ABC transporter membrane protein"
                /gene="BMETHMGA302260c"
CDS             complement(2605079..2606314)
                /note="ABC transporter integral membrane protein"
                /gene="BMETHMGA302261c"
CDS             complement(2606409..2607647)
                /note="extracellular solute-binding protein, family
         I"
                /gene="BMETHMGA302262c"
```

```
CDS             2607764..2608507
                /note="transcriptional regulator, AraC family"
                /gene="BMETHMGA302263"
CDS             complement(2608586..2609026)
                /note="hypothetical protein"
                /gene="BMETHMGA302264c"
CDS             complement(2609045..2609803)
                /note="hypothetical protein"
                /gene="BMETHMGA302265c"
CDS             complement(2609960..2612710)
                /note="Dynamin family protein"
                /gene="BMETHMGA302266c"
CDS             complement(2613101..2613757)
                /note="polar amino acid ABC transproter, innner
       membrane
                subunit"
                /gene="BMETHMGA302267c"
CDS             complement(2613833..2614657)
                /note="extracellular solute-binding protein, family
       3"
                /gene="BMETHMGA302268c"
CDS             complement(2614923..2615711)
                /note="hypothetical protein"
                /gene="BMETHMGA302269c"
CDS             complement(2615708..2616802)
                /note="glycosyltransferase, group 1"
                /gene="BMETHMGA302270c"
CDS             complement(2616842..2617858)
                /note="hypothetical protein"
                /gene="BMETHMGA302271c"
CDS             2618151..2619365
                /note="multidrug resistance protein"
                /gene="BMETHMGA302272"
CDS             2619535..2621157
                /note="glutamyl-tRNA synthetase"
                /gene="BMETHMGA302273"
CDS             complement(2621268..2622329)
                /note="6-phosphogluconolactonase"
                /gene="BMETHMGA302274c"
CDS             complement(2622505..2623362)
                /note="hypothetical protein"
                /gene="BMETHMGA302275c"
CDS             2623673..2624248
                /note="hypothetical protein"
                /gene="BMETHMGA302276"
CDS             2624484..2625746
                /note="ammonium transporter"
                /gene="BMETHMGA302277"
CDS             2625963..2626928
                /note="hypothetical protein"
                /gene="BMETHMGA302278"
CDS             2626930..2627664
                /note="exonuclease RNase T and DNA polymerase III"
```

```
                /gene="BMETHMGA302279"
CDS             2628301..2630310
                /note="carbon starvation protein CstA"
                /gene="BMETHMGA302280"
CDS             2630719..2631120
                /note="cell wall-associated hyrolase containing N-
       terminal
                Lys repeats"
                /gene="BMETHMGA302281"
CDS             2631248..2632207
                /note="D-isomer specific 2-hyroxyacid dehydrogenase"
                /gene="BMETHMGA302282"
CDS             complement(2632910..2633353)
                /note="hypothetical protein"
                /gene="BMETHMGA302284c"
CDS             2633612..2635288
                /note="hypothetical protein"
                /gene="BMETHMGA302285"
CDS             complement(2635378..2636790)
                /note="succinate-semialdehyde dehyrogenase"
                /gene="BMETHMGA302286c"
CDS             complement(2637017..2637745)
                /note="glutamine ABC transporter ATP-binding protein"
                /gene="BMETHMGA302287c"
CDS             complement(2637742..2638386)
                /note="polar amino acid ABC transporter, inner
       membrane
                subunit"
                /gene="BMETHMGA302288c"
CDS             complement(2638448..2639233)
                /note="extracellular solute-binding protein family 3"
                /gene="BMETHMGA302289c"
CDS             2639583..2639879
                /note="glutamate dehydrogenase"
                /gene="BMETHMGA302290"
CDS             2639908..2641320
                /note="glycoside hydrolase family 18"
                /gene="BMETHMGA302291"
CDS             complement(2641370..2643610)
                /note="ATP-dependent DNA helicase RecQ"
                /gene="BMETHMGA302292c"
CDS             complement(2644106..2644477)
                /note="hypothetical protein"
                /gene="BMETHMGA302293c"
CDS             complement(2644495..2644848)
                /note="hypothetical protein"
                /gene="BMETHMGA302294c"
CDS             2646154..2647455
                /note="branched-chain amino acid transporter II
       carrier
                protein"
                /gene="BMETHMGA302295"
CDS             2647847..2648470
```

```
                /note="hypothetical protein"
                /gene="BMETHMGA302296"
CDS             2648467..2649468
                /note="hypothetical protein"
                /gene="BMETHMGA302297"
CDS             2649481..2650269
                /note="LamB/YcsF protein"
                /gene="BMETHMGA302298"
CDS             complement(2650360..2651076)
                /note="hypothetical protein"
                /gene="BMETHMGA302299c"
CDS             2651484..2651855
                /note="hypothetical protein"
                /gene="BMETHMGA302300"
CDS             complement(2651997..2652962)
                /note="coenzyme PQQ synthesis protein"
                /gene="BMETHMGA302301c"
CDS             complement(2654176..2654934)
                /note="phosphanate-transporting ATPase"
                /gene="BMETHMGA302302c"
CDS             complement(2654935..2655744)
                /note="cobalt ABC transporter"
                /gene="BMETHMGA302303c"
CDS             complement(2655808..2656863)
                /note="cobalamin (B12) biosynthesis CbiM protein"
                /gene="BMETHMGA302304c"
CDS             complement(2658107..2659417)
                /note="acetaldehyde dehyrogenase"
                /gene="BMETHMGA302305c"
CDS             complement(2659965..2661275)
                /note="acetaldehyde dehydrogenase"
                /gene="BMETHMGA302306c"
CDS             2661795..2662355
                /note="3D domain-containing protein"
                /gene="BMETHMGA302307"
CDS             complement(2662432..2663340)
                /note="hypothetical protein"
                /gene="BMETHMGA302308c"
CDS             complement(2663461..2664063)
                /note="precorrin-2 dehydrogenase"
                /gene="BMETHMGA302309c"
fasta_record    2664329..2681806
                /note="ConsensusfromContig59 Average coverage:
       505.88"
                /label=ConsensusfromContig59 Average coverage: 505.88
                /colour=10
CDS             complement(2664501..2665610)
                /note="glycosyltransferase, group 1"
                /gene="BMETHMGA302310c"
CDS             complement(2665636..2666094)
                /note="spore coat polysaccharide synthesis
                dTDP-4-deydrorhamnose epimerase"
                /gene="BMETHMGA302311c"
```

```
CDS             complement(2666091..2666951)
                /note="spore coat polysaccharide synthesis
                dTDP-4-dehydroramnose reductase"
                /gene="BMETHMGA302312c"
CDS             complement(2666938..2667765)
                /note="spore coat polysaccharide synthesis protein"
                /gene="BMETHMGA302313c"
CDS             complement(2667894..2668631)
                /note="spore coat polysaccharide synthesis"
                /gene="BMETHMGA302314c"
CDS             2669134..2670684
                /note="glycosyl transferase, group I"
                /gene="BMETHMGA302315"
CDS             2672326..2672949
                /note="cephalosporin hydroxylase"
                /gene="BMETHMGA302316"
CDS             complement(2673888..2674250)
                /note="hypothetical protein"
                /gene="BMETHMGA302317c"
CDS             2674366..2675082
                /note="hypothetical protein"
                /gene="BMETHMGA302318"
CDS             2675882..2677186
                /note="maturation of outermost layer of  the spore"
                /gene="BMETHMGA302319"
CDS             complement(2677961..2678206)
                /note="transposase for insertion sequence element
       OS231B"
                /gene="BMETHMGA302321c"
CDS             complement(2679112..2679768)
                /note="hypothetical protein"
                /gene="BMETHMGA302324c"
CDS             complement(2680143..2681525)
                /note="mannose-6-phosphate isomerase"
                /gene="BMETHMGA302325c"
fasta_record    2681807..2692027
                /note="ConsensusfromContig60 Average coverage:
       492.70"
                /label=ConsensusfromContig60 Average coverage: 492.70
                /colour=11
CDS             2682096..2683577
                /note="carboxy-terminal processing protease"
                /gene="BMETHMGA302326"
CDS             2683792..2684757
                /note="hypothetical protein"
                /gene="BMETHMGA302327"
CDS             2685021..2685356
                /note="hypothetical protein"
                /gene="BMETHMGA302328"
CDS             2685435..2685866
                /note="transcriptional regulator, MarR type"
                /gene="BMETHMGA302329"
CDS             2685952..2686803
```

```
                /note="hypothetical protein"
                /gene="BMETHMGA302330"
CDS             2687587..2688570
                /note="surface antigen"
                /gene="BMETHMGA302332"
CDS             2688567..2689247
                /note="two-component stystem response regulator"
                /gene="BMETHMGA302333"
CDS             2689244..2690644
                /note="two-component histidine kinase"
                /gene="BMETHMGA302334"
CDS             2690692..2691171
                /note="hypothetical protein"
                /gene="BMETHMGA302335"
CDS             2691336..2692091
                /note="hypothetical protein"
                /gene="BMETHMGA302336"
CDS             complement(2692015..2693325)
                /note="hypothetical protein"
                /gene="BMETHMGA302337c"
fasta_record    2692028..2693307
                /note="ConsensusfromContig61 Average coverage:
       1,878.42"
                /label=ConsensusfromContig61 Average coverage:
       1,878.42
                /colour=10
CDS             complement(2693297..2694220)
                /note="ammonium transporter"
                /gene="BMETHMGA302338c"
fasta_record    2693308..2706389
                /note="ConsensusfromContig62 Average coverage:
       513.76"
                /label=ConsensusfromContig62 Average coverage: 513.76
                /colour=11
CDS             complement(2694217..2694585)
                /note="nitrogen regulatory protein P-II"
                /gene="BMETHMGA302339c"
CDS             2694802..2695137
                /note="membrane protein ynfA"
                /gene="BMETHMGA302340"
CDS             2695931..2698048
                /note="ribonucleotide-diphoshpate reductase subunit
       alpha"
                /gene="BMETHMGA302341"
CDS             2698110..2699156
                /note="ribonucleotide-diphoshpate reductase subunit
       beta"
                /gene="BMETHMGA302342"
CDS             2699167..2699601
                /note="flavodoxin"
                /gene="BMETHMGA302343"
CDS             2700283..2700786
                /note="hypothetical protein"
```

```
                     /gene="BMETHMGA302344"
     CDS             complement(2701059..2702891)
                     /note="cadmium efflux P-type ATPase"
                     /gene="BMETHMGA302345c"
     CDS             complement(2704035..2704415)
                     /note="hypothetical protein"
                     /gene="BMETHMGA302348c"
     CDS             complement(2704445..2704792)
                     /note="peptidase domain-containing protein"
                     /gene="BMETHMGA302349c"
     CDS             2704967..2706286
                     /note="Na/H exchanger"
                     /gene="BMETHMGA302350"
     fasta_record    2706390..2717400
                     /note="ConsensusfromContig63 Average coverage:
            590.93"
                     /label=ConsensusfromContig63 Average coverage: 590.93
                     /colour=10
     CDS             complement(2707816..2708190)
                     /note="integrase"
                     /gene="BMETHMGA302352c"
     CDS             2708378..2708905
                     /note="transcriptional modulator of MazE/toxin. MazF"
                     /gene="BMETHMGA302353"
     CDS             complement(2709774..2710229)
                     /note="hypothetical protein"
                     /gene="BMETHMGA302354c"
     CDS             complement(2710343..2711818)
                     /note="type I phosphodiesterase/nucleotide
                     pyrophosphatase"
                     /gene="BMETHMGA302355c"
     CDS             2712494..2713537
                     /note="monogalactosyldiacylglycerol synthase"
                     /gene="BMETHMGA302356"
     CDS             2715729..2716277
                     /note="hypothetical protein"
                     /gene="BMETHMGA302359"
     CDS             2716812..2717186
                     /note="hypothetical protein"
                     /gene="BMETHMGA302360"
     fasta_record    2717401..2721210
                     /note="ConsensusfromContig64 Average coverage:
            872.38"
                     /label=ConsensusfromContig64 Average coverage: 872.38
                     /colour=11
     CDS             2717841..2718575
                     /note="transposase for insertion sequence element
            IS232"
                     /gene="BMETHMGA302361"
     CDS             2718575..2719336
                     /note="ATP-binding protein IStB"
                     /gene="BMETHMGA302362"
     fasta_record    2721211..2724924
```

```
                /note="ConsensusfromContig65 Average coverage:
       557.78"
                /label=ConsensusfromContig65 Average coverage: 557.78
                /colour=10
CDS             complement(2721380..2722039)
                /note="hypothetical protein"
                /gene="BMETHMGA302365c"
CDS             complement(2723285..2723713)
                /note="hypothetical protein"
                /gene="BMETHMGA302366c"
CDS             complement(2724278..2724889)
                /note="cobyrinic acid ac-diamide synthase"
                /gene="BMETHMGA302367c"
fasta_record    2724925..2753197
                /note="ConsensusfromContig66 Average coverage:
       581.56"
                /label=ConsensusfromContig66 Average coverage: 581.56
                /colour=11
CDS             2725614..2727581
                /note="glycogen branching enzyme"
                /gene="BMETHMGA302368"
CDS             2727523..2728668
                /note="glucose-1-phosphate adenylyltransferase"
                /gene="BMETHMGA302369"
CDS             2728833..2729726
                /note="ADP-glucose pyrophosphorylase"
                /gene="BMETHMGA302370"
CDS             2730134..2731174
                /note="glycogen synthase"
                /gene="BMETHMGA302371"
CDS             2731164..2733575
                /note="glycogen phosphorylase"
                /gene="BMETHMGA302372"
CDS             complement(2733647..2734579)
                /note="1,4-dihydroxy-2-naphthoate
       octaprenyltransferase"
                /gene="BMETHMGA302373c"
CDS             2734955..2736400
                /note="isochorismate synthase"
                /gene="BMETHMGA302374"
CDS             2736508..2738151
                /note="2-succinyl-6-hydroxy-2,4-cyclohexadiene-1-
       carboxylate
                synthase"
                /gene="BMETHMGA302375"
CDS             2738151..2738972
                /note="2-succinyl-6-hydroxy-2,4-cyclohexadiene-1-
       carboxylate
                synthase"
                /gene="BMETHMGA302376"
CDS             2739103..2739921
                /note="naphtoate synthase"
                /gene="BMETHMGA302377"
```

```
CDS             2740050..2741528
                /note="O-succinylbenzoic acid-CoA ligase"
                /gene="BMETHMGA302378"
CDS             2742217..2742666
                /note="carbonate dehydratase"
                /gene="BMETHMGA302379"
CDS             2743036..2743509
                /note="S-ribosylhomocysteinase"
                /gene="BMETHMGA302380"
CDS             complement(2743803..2744372)
                /note="spore coat-associated protein"
                /gene="BMETHMGA302381c"
CDS             2745272..2745703
                /note="transcriptional regulator, ArsR family"
                /gene="BMETHMGA302382"
CDS             2746025..2747017
                /note="hypothetical protein"
                /gene="BMETHMGA302383"
CDS             2747092..2747532
                /note="DNA-binding ferritin-like protein"
                /gene="BMETHMGA302384"
CDS             2747652..2748083
                /note="hypothetical protein"
                /gene="BMETHMGA302385"
CDS             2748912..2749271
                /note="hypothetical protein"
                /gene="BMETHMGA302386"
CDS             2749413..2749889
                /note="nucleoside triphosphatase YtkD"
                /gene="BMETHMGA302387"
CDS             complement(2749960..2750859)
                /note="ABC transporter integral membrane protein"
                /gene="BMETHMGA302388c"
CDS             complement(2750750..2751526)
                /note="nitrate/sulfonate/taurine/bicarbonate ABC
                transporter"
                /gene="BMETHMGA302389c"
CDS             complement(2751565..2752557)
                /note="nitrate/sulfonate/taurine/bicarbonate ABC
                transporter"
                /gene="BMETHMGA302390c"
fasta_record    2753198..2753638
                /note="ConsensusfromContig67 Average coverage:
       725.57"
                /label=ConsensusfromContig67 Average coverage: 725.57
                /colour=10
CDS             complement(2753533..2754174)
                /note="hypothetical protein"
                /gene="BMETHMGA302391c"
fasta_record    2753639..2756048
                /note="ConsensusfromContig68 Average coverage:
       528.21"
                /label=ConsensusfromContig68 Average coverage: 528.21
```

```
                 /colour=11
CDS              complement(2754167..2754856)
                 /note="phenylacetate-coenzyme A ligase"
                 /gene="BMETHMGA302392c"
fasta_record     2756049..2779833
                 /note="ConsensusfromContig69 Average coverage:
       524.60"
                 /label=ConsensusfromContig69 Average coverage: 524.60
                 /colour=10
CDS              complement(2756177..2756725)
                 /note="hypothetical protein"
                 /gene="BMETHMGA302394c"
CDS              complement(2756770..2758032)
                 /note="Tetratricopeptide TPR_1 repeat-containing
       protein"
                 /gene="BMETHMGA302395c"
CDS              complement(2758046..2758948)
                 /note="hypothetical protein"
                 /gene="BMETHMGA302396c"
CDS              complement(2759076..2760275)
                 /note="3-phosphoshikimate 1-carboxyvinyltransferase"
                 /gene="BMETHMGA302397c"
CDS              complement(2760393..2761367)
                 /note="Prephenate dehydrogenase"
                 /gene="BMETHMGA302398c"
CDS              complement(2761515..2762609)
                 /note="histidinol-phosphate aminotransferase"
                 /gene="BMETHMGA302399c"
CDS              complement(2762765..2763130)
                 /note="chorismate mutase"
                 /gene="BMETHMGA302400c"
CDS              complement(2763131..2764219)
                 /note="3-dehydroquinate synthase"
                 /gene="BMETHMGA302401c"
CDS              complement(2764219..2765388)
                 /note="chorismate synthase"
                 /gene="BMETHMGA302402c"
CDS              complement(2765631..2766500)
                 /note="chemotactic methyltransferase"
                 /gene="BMETHMGA302403c"
CDS              complement(2766571..2767017)
                 /note="nucleoside diphosphate kinase"
                 /gene="BMETHMGA302404c"
CDS              complement(2767114..2768076)
                 /note="heptaprenyl diphosphate synthase component II"
                 /gene="BMETHMGA302405c"
CDS              complement(2768170..2768880)
                 /note=" 2-heptaprenyl-1,4-naphthoquinone
                 methyltransferase"
                 /gene="BMETHMGA302406c"
CDS              complement(2768888..2769673)
                 /note="heptaprenyl diphosphate synthase component I"
                 /gene="BMETHMGA302407c"
```

```
CDS             complement(2770185..2770751)
                /note="GTP cyclohydrolase I"
                /gene="BMETHMGA302408c"
CDS             complement(2771044..2771316)
                /note="DNA-binding protein HU"
                /gene="BMETHMGA302409c"
CDS             complement(2771896..2773263)
                /note="stage IV sporulation protein A"
                /gene="BMETHMGA302410c"
CDS             complement(2773621..2774319)
                /note="hypothetical protein"
                /gene="BMETHMGA302411c"
CDS             complement(2775006..2775989)
                /note="glycerol-3-phosphate dehydrogenase"
                /gene="BMETHMGA302412c"
CDS             complement(2776073..2777392)
                /note="GTP-binding protein EngA"
                /gene="BMETHMGA302413c"
CDS             complement(2777823..2778719)
                /note="hypothetical protein"
                /gene="BMETHMGA302414c"
CDS             complement(2778716..2779324)
                /note="hypothetical protein"
                /gene="BMETHMGA302415c"
fasta_record    2779834..2791672
                /note="ConsensusfromContig70 Average coverage:
       644.05"
                /label=ConsensusfromContig70 Average coverage: 644.05
                /colour=11
CDS             complement(2779919..2780395)
                /note="ybaK/ebsC protein"
                /gene="BMETHMGA302416c"
CDS             complement(2780524..2781405)
                /note="UDP-glucose pyrophosphorylase"
                /gene="BMETHMGA302417c"
CDS             2781646..2783217
                /note="FAD dependent oxidoreductase"
                /gene="BMETHMGA302418"
CDS             complement(2783323..2784222)
                /note="reductase"
                /gene="BMETHMGA302419c"
CDS             complement(2784343..2784969)
                /note="hypothetical protein"
                /gene="BMETHMGA302420c"
CDS             complement(2784984..2786204)
                /note="hypothetical protein"
                /gene="BMETHMGA302421c"
CDS             2786802..2787893
                /note="peptidase T"
                /gene="BMETHMGA302422"
CDS             2788387..2788926
                /note="hypothetical protein"
                /gene="BMETHMGA302423"
```

```
CDS             complement(2789751..2790077)
                /note="hypothetical protein"
                /gene="BMETHMGA302424c"
CDS             2790438..2790878
                /note="OsmC family protein"
                /gene="BMETHMGA302425"
fasta_record    2791673..2794669
                /note="ConsensusfromContig71 Average coverage:
       540.44"
                /label=ConsensusfromContig71 Average coverage: 540.44
                /colour=10
CDS             2791818..2792420
                /note="aldolase family protein"
                /gene="BMETHMGA302426"
CDS             2793287..2793958
                /note="Transcriptional activator tenA"
                /gene="BMETHMGA302427"
CDS             2793966..2794781
                /note="ABC-type nitrate/sulfonate/bicarbonate
       transport
                system, ATPase  component"
                /gene="BMETHMGA302428"
fasta_record    2794670..2795117
                /note="ConsensusfromContig72 Average coverage:
       1,267.51"
                /label=ConsensusfromContig72 Average coverage:
       1,267.51
                /colour=11
fasta_record    2795118..2795469
                /note="ConsensusfromContig73 Average coverage:
       1,901.96"
                /label=ConsensusfromContig73 Average coverage:
       1,901.96
                /colour=10
fasta_record    2795470..2796564
                /note="ConsensusfromContig74 Average coverage:
       655.92"
                /label=ConsensusfromContig74 Average coverage: 655.92
                /colour=11
fasta_record    2796565..2798612
                /note="ConsensusfromContig75 Average coverage:
       598.96"
                /label=ConsensusfromContig75 Average coverage: 598.96
                /colour=10
CDS             2796732..2797544
                /note="spore germination protein"
                /gene="BMETHMGA302432"
CDS             2797627..2798616
                /note="spore germination protein"
                /gene="BMETHMGA302433"
fasta_record    2798613..2799212
                /note="ConsensusfromContig76 Average coverage:
       543.23"
```

```
                /label=ConsensusfromContig76 Average coverage: 543.23
                /colour=11
CDS             2798751..2799236
                /note="transposase, IS605 OrfB family"
                /gene="BMETHMGA302435"
fasta_record    2799213..2832338
                /note="ConsensusfromContig77 Average coverage:
       524.12"
                /label=ConsensusfromContig77 Average coverage: 524.12
                /colour=10
CDS             complement(2799316..2799837)
                /note="1,2-dihydroxy-3-keto-5-methylthiopentene
                dioxygenase"
                /gene="BMETHMGA302436c"
CDS             complement(2799860..2800495)
                /note="aldolase"
                /gene="BMETHMGA302437c"
CDS             complement(2800492..2801154)
                /note="2-hydroxy-3-keto-5-methylthiopentenyl-1-
       phosphate
                phosphatase"
                /gene="BMETHMGA302438c"
CDS             complement(2801154..2802395)
                /note="2,3-diketo-5-methylthiopentyl-1-phosphate
       enolase"
                /gene="BMETHMGA302439c"
CDS             complement(2802744..2803916)
                /note="transaminase"
                /gene="BMETHMGA302440c"
CDS             2804036..2804818
                /note="Nitrilase/cyanide hydratase and apolipoprotein
                N-acyltransferase"
                /gene="BMETHMGA302441"
CDS             2805085..2806329
                /note="methylthioribose kinase"
                /gene="BMETHMGA302442"
CDS             complement(2806435..2806962)
                /note="methylated-DNA/protein-cysteine
       methyltransferase"
                /gene="BMETHMGA302443c"
CDS             complement(2806969..2808456)
                /note=" two-component sensor histidine kinase"
                /gene="BMETHMGA302444c"
CDS             complement(2808581..2810371)
                /note="methyl-accepting chemotaxis protein"
                /gene="BMETHMGA302445c"
CDS             2810651..2811697
                /note="hypothetical protein"
                /gene="BMETHMGA302446"
CDS             2811748..2813511
                /note="Mg-protoporphyrin IX monomethyl ester
       oxidative
                cyclase"
```

```
                /gene="BMETHMGA302447"
CDS             2813994..2814338
                /note="hypothetical protein"
                /gene="BMETHMGA302448"
CDS             2814400..2814675
                /note="small acid-soluble spore protein"
                /gene="BMETHMGA302449"
CDS             complement(2814696..2815895)
                /note="hypothetical protein"
                /gene="BMETHMGA302450c"
CDS             complement(2815910..2816656)
                /note="RNA polymerase sigma-I factor"
                /gene="BMETHMGA302451c"
CDS             complement(2816977..2817447)
                /note="hypothetical protein"
                /gene="BMETHMGA302452c"
CDS             complement(2817661..2819211)
                /note="aldehyde dehydrogenase"
                /gene="BMETHMGA302453c"
CDS             complement(2819394..2821085)
                /note="peptide chain release factor 3"
                /gene="BMETHMGA302454c"
CDS             complement(2821516..2822136)
                /note="spore cortex-lytic enzyme"
                /gene="BMETHMGA302455c"
CDS             complement(2822725..2823453)
                /note="protein-glutamine gamma-glutamyltransferase"
                /gene="BMETHMGA302456c"
CDS             complement(2823595..2826231)
                /note="cyanophycin synthetase"
                /gene="BMETHMGA302457c"
CDS             complement(2826236..2827039)
                /note="Cyanophycinase"
                /gene="BMETHMGA302458c"
CDS             complement(2827617..2828711)
                /note="spore coat protein CotS"
                /gene="BMETHMGA302459c"
CDS             2828904..2830079
                /note="spore coat protein SA"
                /gene="BMETHMGA302460"
CDS             2830072..2831148
                /note="Spore coat protein S"
                /gene="BMETHMGA302461"
CDS             2831141..2831617
                /note="hypothetical protein"
                /gene="BMETHMGA302462"
CDS             2831695..2832111
                /note="hypothetical protein"
                /gene="BMETHMGA302463"
fasta_record    2832339..2832985
                /note="ConsensusfromContig78 Average coverage:
       1,603.72"
```

```
                 /label=ConsensusfromContig78 Average coverage:
       1,603.72
                 /colour=11
fasta_record     2832986..2833379
                 /note="ConsensusfromContig79 Average coverage:
       1,286.58"
                 /label=ConsensusfromContig79 Average coverage:
       1,286.58
                 /colour=10
fasta_record     2833380..2860142
                 /note="ConsensusfromContig80 Average coverage:
       510.03"
                 /label=ConsensusfromContig80 Average coverage: 510.03
                 /colour=11
CDS              complement(2833676..2836432)
                 /note="hypothetical protein"
                 /gene="BMETHMGA302464c"
CDS              2836691..2836972
                 /note="transposase IS204/IS1001/IS1096/IS1165 family
                 protein"
                 /gene="BMETHMGA302465"
CDS              2837437..2837790
                 /note="transposase IS204/IS1001/IS1096/IS1165 family
                 protein"
                 /gene="BMETHMGA302466"
CDS              complement(2838283..2838729)
                 /note="competence protein CoiA-like family"
                 /gene="BMETHMGA302467c"
CDS              complement(2838826..2842026)
                 /note="helicase domain protein"
                 /gene="BMETHMGA302468c"
CDS              complement(2842067..2842780)
                 /note="conserved hypothetical protein"
                 /gene="BMETHMGA302469c"
CDS              complement(2842784..2844736)
                 /note="PglZ domain family"
                 /gene="BMETHMGA302470c"
CDS              complement(2844847..2847609)
                 /note="DNA methylase N-4/N-6 domain protein"
                 /gene="BMETHMGA302471c"
CDS              complement(2847697..2851404)
                 /note="virulence associated protein"
                 /gene="BMETHMGA302472c"
CDS              complement(2851421..2851861)
                 /note="virulence associated protein"
                 /gene="BMETHMGA302473c"
CDS              complement(2852356..2854743)
                 /note="aconitate hydratase"
                 /gene="BMETHMGA302474c"
CDS              complement(2856262..2857599)
                 /note="L-seryl-tRNA selenium transferase"
                 /gene="BMETHMGA302475c"
CDS              2857779..2858174
```

```
                /note="spore coat protein"
                /gene="BMETHMGA302476"
CDS             2858975..2859790
                /note="selenide, water dikinase"
                /gene="BMETHMGA302477"
fasta_record    2860143..2861500
                /note="ConsensusfromContig81 Average coverage:
       480.91"
                /label=ConsensusfromContig81 Average coverage: 480.91
                /colour=10
CDS             complement(2860255..2861280)
                /note="hypothetical protein"
                /gene="BMETHMGA302478c"
fasta_record    2861501..2867526
                /note="ConsensusfromContig82 Average coverage:
       521.22"
                /label=ConsensusfromContig82 Average coverage: 521.22
                /colour=11
CDS             complement(2861605..2863299)
                /note="oligoendopeptidase F"
                /gene="BMETHMGA302479c"
CDS             complement(2863430..2863891)
                /note="regulatory protein MarR"
                /gene="BMETHMGA302480c"
CDS             2864065..2865639
                /note="hypothetical protein"
                /gene="BMETHMGA302481"
CDS             2865639..2866871
                /note="major facilitator superfamily MFS_1"
                /gene="BMETHMGA302482"
fasta_record    2867527..2872530
                /note="ConsensusfromContig83 Average coverage:
       553.50"
                /label=ConsensusfromContig83 Average coverage: 553.50
                /colour=10
CDS             2870765..2871400
                /note="hypothetical protein"
                /gene="BMETHMGA302485"
fasta_record    2872531..2875560
                /note="ConsensusfromContig84 Average coverage:
       6,303.71"
                /label=ConsensusfromContig84 Average coverage:
       6,303.71
                /colour=11
CDS             2872693..2873010
                /note="hypothetical protein"
                /gene="BMETHMGA302486"
CDS             2874647..2874892
                /note="conserved hypothetical protein"
                /gene="BMETHMGA302488"
fasta_record    2875561..2893439
                /note="ConsensusfromContig85 Average coverage:
       606.27"
```

```
                /label=ConsensusfromContig85 Average coverage: 606.27
                /colour=10
CDS             complement(2875765..2876970)
                /note="PDZ/DHR/GLGF domain protein"
                /gene="BMETHMGA302490c"
CDS             complement(2877297..2877719)
                /note="carboxyl-terminal processing protease"
                /gene="BMETHMGA302491c"
CDS             complement(2877767..2878762)
                /note="carboxyl-terminal processing protease"
                /gene="BMETHMGA302492c"
CDS             complement(2879013..2879693)
                /note="hypothetical protein"
                /gene="BMETHMGA302493c"
CDS             complement(2879975..2881189)
                /note="hypothetical protein"
                /gene="BMETHMGA302494c"
CDS             complement(2881396..2882295)
                /note="cell-division protein"
                /gene="BMETHMGA302495c"
CDS             complement(2882279..2882965)
                /note="cell division ATP-binding protein FtsE"
                /gene="BMETHMGA302496c"
CDS             complement(2883467..2883817)
                /note="hypothetical protein"
                /gene="BMETHMGA302498c"
CDS             complement(2883870..2884733)
                /note="hypothetical protein"
                /gene="BMETHMGA302499c"
CDS             complement(2885068..2886051)
                /note="peptide chain release factor 2"
                /gene="BMETHMGA302500c"
CDS             complement(2886315..2888822)
                /note="preprotein translocase secA subunit"
                /gene="BMETHMGA302501c"
CDS             complement(2889057..2889611)
                /note="hypothetical protein"
                /gene="BMETHMGA302502c"
CDS             complement(2889914..2891074)
                /note="hypothetical protein"
                /gene="BMETHMGA302503c"
CDS             complement(2891238..2891771)
                /note="Chromate transporter"
                /gene="BMETHMGA302504c"
CDS             complement(2891768..2892367)
                /note="chromate transporter"
                /gene="BMETHMGA302505c"
CDS             complement(2892583..2893137)
                /note="ribosomal subunit interface protein "
                /gene="BMETHMGA302506c"
fasta_record    2893440..2907719
                /note="ConsensusfromContig86 Average coverage:
       628.09"
```

```
                     /label=ConsensusfromContig86 Average coverage: 628.09
                     /colour=11
CDS                  2893705..2895225
                     /note=" hypothetical protein"
                     /gene="BMETHMGA302507"
CDS                  2895222..2895941
                     /note="phosphoribosyltransferase"
                     /gene="BMETHMGA302508"
CDS                  2896023..2896457
                     /note="hypothetical protein"
                     /gene="BMETHMGA302509"
CDS                  2896894..2897376
                     /note="FlgN family protein"
                     /gene="BMETHMGA302510"
CDS                  2897395..2899002
                     /note="flagellar hook-associated protein K"
                     /gene="BMETHMGA302511"
CDS                  2899016..2899879
                     /note="flagellar hook-associated protein L"
                     /gene="BMETHMGA302512"
CDS                  2899993..2900559
                     /note="hypothetical protein"
                     /gene="BMETHMGA302513"
CDS                  2900572..2901021
                     /note="flagellar assembly protein FliW"
                     /gene="BMETHMGA302514"
CDS                  2901398..2903011
                     /note="flagellin domain-containing protein"
                     /gene="BMETHMGA302515"
CDS                  complement(2904042..2904755)
                     /note="LAGLIDADG DNA endonuclease"
                     /gene="BMETHMGA302516c"
CDS                  2904852..2905232
                     /note="flagellar protein FlaG"
                     /gene="BMETHMGA302517"
CDS                  2905241..2906827
                     /note="flagellar capping protein"
                     /gene="BMETHMGA302518"
CDS                  2906900..2907301
                     /note="flagellar protein FliS"
                     /gene="BMETHMGA302519"
fasta_record         2907720..2921080
                     /note="ConsensusfromContig87 Average coverage:
         567.13"
                     /label=ConsensusfromContig87 Average coverage: 567.13
                     /colour=10
CDS                  complement(2907832..2908734)
                     /note="hypothetical protein"
                     /gene="BMETHMGA302521c"
CDS                  complement(2908718..2909512)
                     /note="hypothetical protein"
                     /gene="BMETHMGA302522c"
CDS                  complement(2909615..2910661)
```

```
                /note="hypothetical protein"
                /gene="BMETHMGA302523c"
CDS             complement(2910579..2912408)
                /note="morphogenetic protein associated with SpoVID"
                /gene="BMETHMGA302524c"
CDS             complement(2912573..2913643)
                /note="Quinolinate synthetase A"
                /gene="BMETHMGA302525c"
CDS             complement(2913692..2914597)
                /note="nicotinate-nucleotide pyrophosphorylase"
                /gene="BMETHMGA302526c"
CDS             complement(2914497..2916086)
                /note="L-aspartate oxidase"
                /gene="BMETHMGA302527c"
CDS             2916195..2917334
                /note="cysteine desulfurase"
                /gene="BMETHMGA302528"
CDS             complement(2917386..2918300)
                /note="prephenate dehydratase"
                /gene="BMETHMGA302529c"
CDS             complement(2918518..2918973)
                /note="amino acid-binding ACT domain protein"
                /gene="BMETHMGA302530c"
CDS             complement(2918997..2920286)
                /note="GTPase"
                /gene="BMETHMGA302531c"
CDS             complement(2920349..2920885)
                /note="sporulation initiation phosphotransferase B"
                /gene="BMETHMGA302532c"
fasta_record    2921081..3024672
                /note="ConsensusfromContig88 Average coverage:
       566.92"
                /label=ConsensusfromContig88 Average coverage: 566.92
                /colour=11
CDS             2921455..2922006
                /note="translation initiation factor IF-3"
                /gene="BMETHMGA302533"
CDS             2922255..2922614
                /note="50S ribosomal protein L20"
                /gene="BMETHMGA302534"
CDS             2922765..2923337
                /note="hypothetical protein"
                /gene="BMETHMGA302535"
CDS             complement(2923393..2923797)
                /note="hypothetical protein"
                /gene="BMETHMGA302537c"
CDS             2923934..2924422
                /note="dUTPase"
                /gene="BMETHMGA302538"
CDS             2924397..2925581
                /note="glucanase/ deblocking aminopeptidase"
                /gene="BMETHMGA302539"
CDS             2926153..2926695
```

```
                     /note="23S rRNA methyltransferase"
                     /gene="BMETHMGA302541"
     CDS             2927083..2928120
                     /note="phenylalanyl-tRNA synthetase alpha subunit"
                     /gene="BMETHMGA302542"
     CDS             2928142..2930556
                     /note="phenylalanyl-tRNA synthetase beta subunit"
                     /gene="BMETHMGA302543"
     CDS             complement(2930670..2931443)
                     /note="ribonuclease HIII"
                     /gene="BMETHMGA302544c"
     CDS             2932075..2932617
                     /note="colicin V production protein"
                     /gene="BMETHMGA302545"
     CDS             2932692..2934422
                     /note="DNA-directed DNA polymerase X"
                     /gene="BMETHMGA302546"
     CDS             2934446..2935813
                     /note="recombination and DNA strand exchange
          inhibitor
                     protein"
                     /gene="BMETHMGA302547"
     CDS             2936009..2936803
                     /note="recombination and DNA strand exchange
          inhibitor
                     protein"
                     /gene="BMETHMGA302548"
     CDS             2936822..2937226
                     /note="hypothetical protein"
                     /gene="BMETHMGA302549"
     CDS             2937525..2939228
                     /note="acyl-CoA synthase"
                     /gene="BMETHMGA302550"
     CDS             2939452..2940048
                     /note="transcriptional regulator (TetR/AcrR family)
                     protein"
                     /gene="BMETHMGA302551"
     CDS             2940105..2940848
                     /note="enoyl-CoA hydratase"
                     /gene="BMETHMGA302552"
     CDS             2940974..2941747
                     /note="electron transfer flavoprotein beta subunit"
                     /gene="BMETHMGA302553"
     CDS             2941767..2942762
                     /note="electron transfer flavoprotein, alpha subunit"
                     /gene="BMETHMGA302554"
     CDS             2942946..2943260
                     /note="thioredoxin"
                     /gene="BMETHMGA302555"
     CDS             2943441..2945213
                     /note="excinuclease ABC subunit C"
                     /gene="BMETHMGA302556"
     CDS             2945741..2946850
```

```
                /note="Aspartokinase 2"
                /gene="BMETHMGA302557"
CDS             complement(2947005..2947460)
                /note="hypothetical protein "
                /gene="BMETHMGA302558c"
CDS             2947712..2948440
                /note="succinate dehydrogenase"
                /gene="BMETHMGA302559"
CDS             2948497..2950257
                /note="succinate dehydrogenase"
                /gene="BMETHMGA302560"
CDS             2950261..2951028
                /note="succinate dehydrogenase iron-sulfur subunit"
                /gene="BMETHMGA302561"
CDS             complement(2951641..2953053)
                /note="short chain fatty acid transporter"
                /gene="BMETHMGA302562c"
CDS             2953296..2953754
                /note="transcriptional regulator, MarR family"
                /gene="BMETHMGA302563"
CDS             2953828..2954562
                /note="glutamate racemase"
                /gene="BMETHMGA302564"
CDS             2954803..2955876
                /note="spore germination protein GerM"
                /gene="BMETHMGA302565"
CDS             2956009..2956779
                /note="ribonuclease PH"
                /gene="BMETHMGA302566"
CDS             2956837..2957391
                /note="nucleoside-triphosphatase"
                /gene="BMETHMGA302567"
CDS             2957388..2957906
                /note="phosphodiesterase"
                /gene="BMETHMGA302568"
CDS             2958475..2959284
                /note="integral membrane protein"
                /gene="BMETHMGA302569"
CDS             2960123..2960890
                /note="branched-chain amino acid aminotransferase"
                /gene="BMETHMGA302570"
CDS             2961239..2962954
                /note="acetolactate synthase, large subunit"
                /gene="BMETHMGA302571"
CDS             2962951..2963469
                /note="acetolactate synthase, small subunit"
                /gene="BMETHMGA302572"
CDS             2963505..2964527
                /note="ketol-acid reductoisomerase"
                /gene="BMETHMGA302573"
CDS             2964742..2966052
                /note="2-isopropylmalate synthase"
                /gene="BMETHMGA302574"
```

```
CDS             2966052..2967176
                /note="3-isopropylmalate dehydrogenase"
                /gene="BMETHMGA302575"
CDS             2967304..2968770
                /note="3-isopropylmalate dehydratase, large subunit"
                /gene="BMETHMGA302576"
CDS             2968786..2969376
                /note="3-isopropylmalate dehydratase, small subunit"
                /gene="BMETHMGA302577"
CDS             2969462..2970487
                /note="putative hydrolase"
                /gene="BMETHMGA302578"
CDS             2970657..2971967
                /note="trigger factor"
                /gene="BMETHMGA302579"
CDS             2972284..2973549
                /note="ATP-dependent Clp protease, ATP-binding
       subunit
                ClpX"
                /gene="BMETHMGA302580"
CDS             2973834..2975501
                /note="ATP-dependent protease"
                /gene="BMETHMGA302581"
CDS             2975646..2977973
                /note="ATP-dependent protease La"
                /gene="BMETHMGA302582"
CDS             2977970..2978551
                /note="GTP-binding protein"
                /gene="BMETHMGA302583"
CDS             complement(2978598..2978966)
                /note="hypothetical protein"
                /gene="BMETHMGA302584c"
CDS             2979342..2980688
                /note="glutamyl-tRNA reductase"
                /gene="BMETHMGA302585"
CDS             2980703..2981536
                /note="cytochrome c assembly protein"
                /gene="BMETHMGA302586"
CDS             2981552..2982487
                /note="porphobilinogen deaminase"
                /gene="BMETHMGA302587"
CDS             2982484..2983257
                /note="Uroporphyrinogen III synthase"
                /gene="BMETHMGA302588"
CDS             2983274..2984251
                /note="Porphobilinogen synthase"
                /gene="BMETHMGA302589"
CDS             2984273..2985562
                /note="glutamate-1-semialdehyde aminotransferase"
                /gene="BMETHMGA302590"
CDS             2985729..2986946
                /note="stage VI sporulation protein D"
                /gene="BMETHMGA302591"
```

```
CDS             2987102..2988136
                /note="spore coat protein YsxE"
                /gene="BMETHMGA302592"
CDS             2988747..2991443
                /note="valyl-tRNA synthetase"
                /gene="BMETHMGA302593"
CDS             2991453..2992841
                /note="folyl-polyglutamate synthetase"
                /gene="BMETHMGA302594"
CDS             2992927..2994639
                /note="two-component sensor histidine kinase"
                /gene="BMETHMGA302595"
CDS             2995049..2997658
                /note="hypothetical protein"
                /gene="BMETHMGA302596"
CDS             complement(2997697..2999841)
                /note="hypothetical protein"
                /gene="BMETHMGA302597c"
CDS             complement(2999917..3000426)
                /note="hypothetical protein"
                /gene="BMETHMGA302598c"
CDS             complement(3000473..3000895)
                /note="hypothetical protein"
                /gene="BMETHMGA302599c"
CDS             3001150..3002502
                /note="hypothetical protein"
                /gene="BMETHMGA302600"
CDS             3002546..3003331
                /note="hypothetical protein"
                /gene="BMETHMGA302601"
CDS             3003608..3004717
                /note="hypothetical protein"
                /gene="BMETHMGA302602"
CDS             3004741..3006402
                /note="ATPase involved in pili biogenesis"
                /gene="BMETHMGA302603"
CDS             3006418..3007458
                /note="twitching motility protein"
                /gene="BMETHMGA302604"
CDS             3007460..3008665
                /note="Type II secretion system F domain protein"
                /gene="BMETHMGA302605"
CDS             3008755..3009207
                /note="Type II secretory pathway pseudopilin PulG"
                /gene="BMETHMGA302606"
CDS             3009216..3009968
                /note="signal peptidase"
                /gene="BMETHMGA302607"
CDS             3009981..3010964
                /note="hypothetical protein"
                /gene="BMETHMGA302608"
CDS             3010967..3011641
                /note="Fimbrial assembly family protein"
```

```
                     /gene="BMETHMGA302609"
     CDS             3011638..3012441
                     /note="hypothetical protein"
                     /gene="BMETHMGA302610"
     CDS             3012398..3012889
                     /note="Pilus assembly protein PilE"
                     /gene="BMETHMGA302611"
     CDS             3012986..3014050
                     /note="hypothetical protein"
                     /gene="BMETHMGA302612"
     CDS             3014446..3015018
                     /note="Septum formation protein Maf"
                     /gene="BMETHMGA302613"
     CDS             3015065..3015772
                     /note="DNA repair protein RadC"
                     /gene="BMETHMGA302614"
     CDS             3015929..3016951
                     /note="rod shape-determining protein MreB"
                     /gene="BMETHMGA302615"
     CDS             3016976..3017860
                     /note="rod shape-determining protein MreC"
                     /gene="BMETHMGA302616"
     CDS             3017881..3018372
                     /note="rod shape-determining protein MreD"
                     /gene="BMETHMGA302617"
     CDS             3018622..3019302
                     /note="septum formation inhibitor"
                     /gene="BMETHMGA302618"
     CDS             3019304..3020110
                     /note="septum site-determining protein MinD"
                     /gene="BMETHMGA302619"
     CDS             3020226..3020999
                     /note="inhibitor of SpoIVFB"
                     /gene="BMETHMGA302620"
     CDS             3021124..3021864
                     /note="peptidase M50"
                     /gene="BMETHMGA302621"
     CDS             3021897..3023366
                     /note="ribonuclease, Rne/Rng family"
                     /gene="BMETHMGA302622"
     CDS             3023601..3023909
                     /note="50S ribosomal protein L21"
                     /gene="BMETHMGA302623"
     CDS             3023922..3024251
                     /note="hypothetical protein"
                     /gene="BMETHMGA302624"
     CDS             3024267..3024557
                     /note="50S ribosomal protein L27"
                     /gene="BMETHMGA302625"
     fasta_record    3024673..3084824
                     /note="ConsensusfromContig89 Average coverage:
              536.99"
                     /label=ConsensusfromContig89 Average coverage: 536.99
```

```
                /colour=10
CDS             3025222..3025740
                /note="YxaI protein"
                /gene="BMETHMGA302626"
CDS             3026192..3027268
                /note="putative efflux transporter"
                /gene="BMETHMGA302627"
CDS             3027341..3028246
                /note="modulation of CheA activity in response to
                attractants"
                /gene="BMETHMGA302628"
CDS             complement(3028278..3028793)
                /note="hypothetical protein"
                /gene="BMETHMGA302629c"
CDS             3029132..3030859
                /note="diguanylate cyclase/phosphodiesterase"
                /gene="BMETHMGA302630"
CDS             complement(3030883..3031755)
                /note="metallophosphoesterase"
                /gene="BMETHMGA302631c"
CDS             complement(3031978..3032742)
                /note="metallophosphoesterase"
                /gene="BMETHMGA302632c"
CDS             3032915..3034150
                /note="diguanylate phosphodiesterase"
                /gene="BMETHMGA302633"
CDS             complement(3034664..3034972)
                /note="Glutaredoxin-like protein, YruB-family"
                /gene="BMETHMGA302634c"
CDS             complement(3034933..3035925)
                /note="metallo-beta-lactamase family protein"
                /gene="BMETHMGA302635c"
CDS             3036401..3036865
                /note="CBS domain-containing protein"
                /gene="BMETHMGA302636"
CDS             3037032..3038477
                /note="drug resistance transporter, EmrB/QacA
        subfamily"
                /gene="BMETHMGA302637"
CDS             3038529..3039422
                /note="transcriptional repressor of citB and citZ"
                /gene="BMETHMGA302638"
CDS             3039532..3040242
                /note="Tetrahydrodipicolinate N-succinyltransferase"
                /gene="BMETHMGA302639"
CDS             3040318..3041442
                /note="hippurate hydrolase"
                /gene="BMETHMGA302640"
CDS             3041855..3042403
                /note="Alkyl hydroperoxide reductase"
                /gene="BMETHMGA302641"
CDS             3042499..3042972
                /note="Thiol-disulfide oxidoreductase ykuV"
```

```
                 /gene="BMETHMGA302642"
CDS              3043079..3043987
                 /note="hypothetical protein"
                 /gene="BMETHMGA302643"
CDS              complement(3044020..3045687)
                 /note="ribonuclease J 1"
                 /gene="BMETHMGA302644c"
CDS              complement(3046387..3046941)
                 /note="peptide deformylase 2 "
                 /gene="BMETHMGA302645c"
CDS              3047978..3049093
                 /note="pyruvate dehydrogenase"
                 /gene="BMETHMGA302646"
CDS              3049097..3050074
                 /note="pyruvate dehydrogenase E1 component subunit
       beta"
                 /gene="BMETHMGA302647"
CDS              3050183..3051541
                 /note="branched-chain alpha-keto acid dehydrogenase
                 subunit E2"
                 /gene="BMETHMGA302648"
CDS              3051546..3052955
                 /note="dihydrolipoamide dehydrogenase"
                 /gene="BMETHMGA302649"
CDS              complement(3053531..3053968)
                 /note="hypothetical protein"
                 /gene="BMETHMGA302650c"
CDS              3054119..3054418
                 /note="hypothetical protein"
                 /gene="BMETHMGA302651"
CDS              complement(3055114..3056301)
                 /note="arginine decarboxylase"
                 /gene="BMETHMGA302652c"
CDS              3056796..3057764
                 /note="2-nitropropane dioxygenase"
                 /gene="BMETHMGA302653"
CDS              3057822..3058106
                 /note="hypothetical protein"
                 /gene="BMETHMGA302654"
CDS              3058736..3059575
                 /note="inositol monophosphatase"
                 /gene="BMETHMGA302656"
CDS              complement(3059796..3061094)
                 /note="Manganese transport protein mntH"
                 /gene="BMETHMGA302657c"
CDS              3061869..3063581
                 /note="GTP-binding protein TypA"
                 /gene="BMETHMGA302658"
CDS              3063600..3063914
                 /note="hypothetical protein"
                 /gene="BMETHMGA302659"
CDS              3063922..3064212
                 /note="hypothetical protein"
```

```
                     /gene="BMETHMGA302660"
CDS                  3064730..3065218
                     /note="pyridoxamine 5'-phosphate oxidase-related
                     FMN-binding protein"
                     /gene="BMETHMGA302661"
CDS                  complement(3065278..3065856)
                     /note="sporulation lipoprotein, YhcN/YlaJ family"
                     /gene="BMETHMGA302662c"
CDS                  3066217..3067416
                     /note="phosphate starvation-induced protein"
                     /gene="BMETHMGA302663"
CDS                  complement(3067520..3068017)
                     /note="hypothetical protein"
                     /gene="BMETHMGA302664c"
CDS                  3068247..3068531
                     /note="hypothetical protein"
                     /gene="BMETHMGA302665"
CDS                  3068790..3069956
                     /note="cell cycle protein"
                     /gene="BMETHMGA302666"
CDS                  3070053..3073496
                     /note="pyruvate carboxylase"
                     /gene="BMETHMGA302667"
CDS                  complement(3073570..3074505)
                     /note="heme O oxygenase"
                     /gene="BMETHMGA302668c"
CDS                  3075210..3076148
                     /note="protoheme IX farnesyltransferase"
                     /gene="BMETHMGA302669"
CDS                  3076268..3077338
                     /note="cytochrome c oxidase, subunit II"
                     /gene="BMETHMGA302670"
CDS                  3077573..3079237
                     /note="cytochrome c oxidase subunit I"
                     /gene="BMETHMGA302671"
CDS                  3079237..3079866
                     /note="cytochrome c oxidase subunit III"
                     /gene="BMETHMGA302672"
CDS                  3079871..3080206
                     /note="cytochrome c oxidase, subunit IV"
                     /gene="BMETHMGA302673"
CDS                  3080552..3081250
                     /note="cytochrome c oxidase assembly factor CtaG"
                     /gene="BMETHMGA302674"
CDS                  3081545..3082009
                     /note="hypothetical protein"
                     /gene="BMETHMGA302675"
CDS                  complement(3082107..3083165)
                     /note="hypothetical protein"
                     /gene="BMETHMGA302676c"
CDS                  complement(3083418..3083780)
                     /note="conserved hypothetical protein"
                     /gene="BMETHMGA302677c"
```

```
CDS             3084281..3084712
                /note="CBS domain containing protein"
                /gene="BMETHMGA302678"
fasta_record    3084825..3106553
                /note="ConsensusfromContig90 Average coverage:
       492.06"
                /label=ConsensusfromContig90 Average coverage: 492.06
                /colour=11
CDS             3084831..3086264
                /note="2-oxoglutarate dehydrogenase E1 component"
                /gene="BMETHMGA302679"
CDS             3086245..3087534
                /note="2-oxoglutarate dehydrogenase E2 component"
                /gene="BMETHMGA302680"
CDS             3087947..3088735
                /note="hypothetical protein"
                /gene="BMETHMGA302681"
CDS             3088680..3089153
                /note="hypothetical protein"
                /gene="BMETHMGA302682"
CDS             3089594..3090013
                /note="acyl-CoA synthase"
                /gene="BMETHMGA302684"
CDS             3090010..3090453
                /note="Long-chain-fatty-acid--CoA ligase"
                /gene="BMETHMGA302685"
CDS             3091223..3092566
                /note="hemolysin"
                /gene="BMETHMGA302686"
CDS             complement(3092661..3093191)
                /note="hypothetical protein"
                /gene="BMETHMGA302687c"
CDS             3093366..3093785
                /note="CrcB protein"
                /gene="BMETHMGA302688"
CDS             3093782..3094135
                /note="CrcB protein"
                /gene="BMETHMGA302689"
CDS             complement(3094235..3095008)
                /note="OxaA-like protein precursor "
                /gene="BMETHMGA302690c"
CDS             complement(3095532..3096125)
                /note="biotin synthase"
                /gene="BMETHMGA302692c"
CDS             complement(3096364..3097071)
                /note="pyruvate formate-lyase activating enzyme"
                /gene="BMETHMGA302693c"
CDS             complement(3097145..3099370)
                /note="Pyruvate-formate lyase"
                /gene="BMETHMGA302694c"
CDS             complement(3099811..3100230)
                /note="hypothetical protein"
                /gene="BMETHMGA302695c"
```

```
CDS             3100476..3101402
                /note="hypothetical protein"
                /gene="BMETHMGA302696"
CDS             3101827..3102903
                /note=" NLP/P60 protein"
                /gene="BMETHMGA302697"
CDS             complement(3103024..3103677)
                /note="molybdate ABC transporter, inner membrane
       subunit"
                /gene="BMETHMGA302698c"
CDS             complement(3104383..3106002)
                /note="ABC efflux transporter ATP-binding protein"
                /gene="BMETHMGA302699c"
fasta_record    3106554..3141638
                /note="ConsensusfromContig91 Average coverage:
       629.61"
                /label=ConsensusfromContig91 Average coverage: 629.61
                /colour=10
CDS             3106828..3107538
                /note="DNA-binding response regulator"
                /gene="BMETHMGA302700"
CDS             3107548..3109368
                /note="PAS sensor protein"
                /gene="BMETHMGA302701"
CDS             3109359..3110696
                /note="regulator of YycFG"
                /gene="BMETHMGA302702"
CDS             3110917..3111477
                /note="hypothetical protein"
                /gene="BMETHMGA302703"
CDS             3111485..3112279
                /note="metal-dependent hydrolase of the beta-
       lactamase
                superfamily III"
                /gene="BMETHMGA302704"
CDS             3112819..3113961
                /note="serine protease Do"
                /gene="BMETHMGA302705"
CDS             3114428..3114907
                /note="ribosomal RNA large subunit methyltransferase
       H"
                /gene="BMETHMGA302706"
CDS             complement(3115079..3116224)
                /note="Monogalactosyldiacylglycerol synthase"
                /gene="BMETHMGA302707c"
CDS             3116719..3117873
                /note="major facilitator superfamily MFS_1"
                /gene="BMETHMGA302708"
CDS             complement(3118045..3119298)
                /note="integral membrane protein"
                /gene="BMETHMGA302709c"
CDS             complement(3119571..3120158)
                /note="hypothetical protein"
```

```
                     /gene="BMETHMGA302710c"
     CDS             3120731..3122197
                     /note="glycosyl transferase family protein"
                     /gene="BMETHMGA302711"
     CDS             3122457..3122909
                     /note="OsmC family protein"
                     /gene="BMETHMGA302712"
     CDS             3123512..3124900
                     /note="family 1 extracellular solute-binding protein"
                     /gene="BMETHMGA302713"
     CDS             3125000..3125857
                     /note="ABC transporter integral membrane protein"
                     /gene="BMETHMGA302714"
     CDS             3125874..3126770
                     /note="ABC transporter integral membrane protein"
                     /gene="BMETHMGA302715"
     CDS             3126767..3127876
                     /note="glycerol-3-phosphate-transporting ATPase"
                     /gene="BMETHMGA302716"
     CDS             3128347..3129000
                     /note="transcriptional regulator, TetR family"
                     /gene="BMETHMGA302717"
     CDS             3129036..3129920
                     /note="ABC transporter related protein"
                     /gene="BMETHMGA302718"
     CDS             3129926..3130723
                     /note="putative ABC-2 type transport system permease
                     protein"
                     /gene="BMETHMGA302719"
     CDS             3130741..3131538
                     /note="ABC-2 type transporter "
                     /gene="BMETHMGA302720"
     CDS             complement(3131653..3132357)
                     /note="hypothetical protein"
                     /gene="BMETHMGA302721c"
     CDS             complement(3132368..3133153)
                     /note="ferrichrome ABC transporter ATP-binding
            protein"
                     /gene="BMETHMGA302722c"
     CDS             complement(3133166..3134173)
                     /note="ferrichrome ABC transporter permease"
                     /gene="BMETHMGA302723c"
     CDS             complement(3134170..3134934)
                     /note="ferrichrome ABC transporter permease"
                     /gene="BMETHMGA302724c"
     CDS             3135449..3136399
                     /note="ferrichrome ABC transporter ferrichrome-
            binding
                     protein"
                     /gene="BMETHMGA302725"
     CDS             complement(3136486..3137535)
                     /note="oenzyme F420-dependent N5,N10-methylene
                     tetrahydromethanopterin  reductase"
```

```
                     /gene="BMETHMGA302726c"
     CDS             complement(3137659..3138336)
                     /note="NAD-dependent epimerase/dehydratase"
                     /gene="BMETHMGA302727c"
     CDS             complement(3138456..3139133)
                     /note="iron dependent repressor"
                     /gene="BMETHMGA302728c"
     CDS             3139528..3139827
                     /note="hypothetical protein"
                     /gene="BMETHMGA302729"
     CDS             3139858..3141534
                     /note="2-polyprenylphenol 6-hydroxylase"
                     /gene="BMETHMGA302730"
     fasta_record    3141639..3147164
                     /note="ConsensusfromContig92 Average coverage:
            615.86"
                     /label=ConsensusfromContig92 Average coverage: 615.86
                     /colour=11
     CDS             complement(3141923..3144046)
                     /note="PTS system, glucose-specific IIBC subunit"
                     /gene="BMETHMGA302731c"
     CDS             complement(3144161..3145030)
                     /note="transcriptional antiterminator, BglG family
                     protein"
                     /gene="BMETHMGA302732c"
     CDS             complement(3145159..3146493)
                     /note="Xanthine/uracil/vitamin C permease"
                     /gene="BMETHMGA302733c"
     fasta_record    3147165..3156830
                     /note="ConsensusfromContig93 Average coverage:
            522.21"
                     /label=ConsensusfromContig93 Average coverage: 522.21
                     /colour=10
     CDS             complement(3147224..3148378)
                     /note="Transporter, MFS superfamily"
                     /gene="BMETHMGA302734c"
     CDS             complement(3149023..3149550)
                     /note="hypothetical protein"
                     /gene="BMETHMGA302735c"
     CDS             complement(3149781..3150575)
                     /note="hypothetical protein"
                     /gene="BMETHMGA302736c"
     CDS             3151913..3153070
                     /note="NAD-dependent methanol dehydrogenase"
                     /gene="BMETHMGA302737"
     CDS             3153333..3153848
                     /note="DNA protection during starvation protein"
                     /gene="BMETHMGA302738"
     CDS             complement(3154390..3155589)
                     /note=" Oxalate decarboxylase oxdD"
                     /gene="BMETHMGA302739c"
     CDS             3155756..3156577
                     /note="Zn-dependent hydrolase"
```

```
                     /gene="BMETHMGA302740"
fasta_record         3156831..3193151
                     /note="ConsensusfromContig94 Average coverage:
         575.98"
                     /label=ConsensusfromContig94 Average coverage: 575.98
                     /colour=11
CDS                  complement(3156970..3158388)
                     /note="hypothetical protein"
                     /gene="BMETHMGA302741c"
CDS                  complement(3158404..3159981)
                     /note="hypothetical protein"
                     /gene="BMETHMGA302742c"
CDS                  complement(3160006..3160491)
                     /note="hypothetical protein"
                     /gene="BMETHMGA302743c"
CDS                  complement(3160582..3161196)
                     /note="hypothetical protein"
                     /gene="BMETHMGA302744c"
CDS                  3161678..3162181
                     /note="ComK family protein"
                     /gene="BMETHMGA302745"
CDS                  3162654..3164387
                     /note="methyl-accepting chemotaxis protein"
                     /gene="BMETHMGA302746"
CDS                  3164742..3165230
                     /note="RNA polymerase sigma factor, sigma-70 family"
                     /gene="BMETHMGA302747"
CDS                  3165452..3166114
                     /note="hypothetical protein"
                     /gene="BMETHMGA302748"
CDS                  complement(3166192..3168018)
                     /note="diguanylate cyclase"
                     /gene="BMETHMGA302749c"
CDS                  complement(3168186..3168785)
                     /note="hypothetical protein"
                     /gene="BMETHMGA302750c"
CDS                  3169116..3169571
                     /note="hypothetical protein"
                     /gene="BMETHMGA302751"
CDS                  3169746..3171080
                     /note="UDP-glucose 6-dehydrogenase"
                     /gene="BMETHMGA302752"
CDS                  3171173..3171835
                     /note="hypothetical protein"
                     /gene="BMETHMGA302753"
CDS                  3172135..3173061
                     /note="hypothetical protein"
                     /gene="BMETHMGA302754"
CDS                  3173194..3176682
                     /note="alkaline serine protease, subtilase family
         protein"
                     /gene="BMETHMGA302755"
CDS                  3176789..3177334
```

```
                     /note="hypothetical protein"
                     /gene="BMETHMGA302756"
     CDS             complement(3177368..3179476)
                     /note="hypothetical protein"
                     /gene="BMETHMGA302757c"
     CDS             complement(3179583..3181175)
                     /note="hypothetical protein"
                     /gene="BMETHMGA302758c"
     CDS             complement(3181185..3182639)
                     /note="alginate O-acetyltransferase"
                     /gene="BMETHMGA302759c"
     CDS             3182804..3184048
                     /note="hypothetical protein"
                     /gene="BMETHMGA302760"
     CDS             3184096..3184995
                     /note="hypothetical protein"
                     /gene="BMETHMGA302761"
     CDS             3185292..3187757
                     /note="hypothetical protein"
                     /gene="BMETHMGA302762"
     CDS             3188052..3188459
                     /note="hypothetical protein"
                     /gene="BMETHMGA302763"
     CDS             3188735..3189436
                     /note="hypothetical protein"
                     /gene="BMETHMGA302764"
     CDS             complement(3189605..3190492)
                     /note="hypothetical protein"
                     /gene="BMETHMGA302765c"
     CDS             complement(3190505..3192877)
                     /note="translocase subunit secA 2"
                     /gene="BMETHMGA302766c"
     fasta_record    3193152..3215213
                     /note="ConsensusfromContig95 Average coverage:
            552.51"
                     /label=ConsensusfromContig95 Average coverage: 552.51
                     /colour=10
     CDS             complement(3193251..3193655)
                     /note="diacylglycerol kinase"
                     /gene="BMETHMGA302767c"
     CDS             complement(3193633..3194103)
                     /note="Putative metalloprotease"
                     /gene="BMETHMGA302768c"
     CDS             complement(3194100..3196313)
                     /note="metal dependent phosphohydrolase"
                     /gene="BMETHMGA302769c"
     CDS             complement(3196433..3197398)
                     /note="phosphate starvation-induced protein PhoH"
                     /gene="BMETHMGA302770c"
     CDS             complement(3197401..3198588)
                     /note="stage IV sporulation protein"
                     /gene="BMETHMGA302771c"
     CDS             complement(3198604..3198909)
```

```
                     /note="sporulation protein YqfC"
                     /gene="BMETHMGA302772c"
     CDS             complement(3199081..3199563)
                     /note="hypothetical protein"
                     /gene="BMETHMGA302773c"
     CDS             complement(3199606..3200604)
                     /note="seryl-tRNA synthetase"
                     /gene="BMETHMGA302774c"
     CDS             complement(3200620..3201897)
                     /note="putative membrane bound hydrolase"
                     /gene="BMETHMGA302775c"
     CDS             complement(3202150..3202596)
                     /note="hypothetical protein"
                     /gene="BMETHMGA302776c"
     CDS             complement(3202909..3203574)
                     /note="deoxyribose phosphate aldolase"
                     /gene="BMETHMGA302777c"
     CDS             complement(3203758..3205110)
                     /note="ribosomal protein S12 methylthiotransferase"
                     /gene="BMETHMGA302778c"
     CDS             complement(3205119..3205784)
                     /note="16S ribosomal RNA methyltransferase RsmE"
                     /gene="BMETHMGA302779c"
     CDS             complement(3205963..3206922)
                     /note="50S ribosomal protein L11 methyltransferase"
                     /gene="BMETHMGA302780c"
     CDS             complement(3206922..3208049)
                     /note="chaperone protein DnaJ"
                     /gene="BMETHMGA302781c"
     CDS             complement(3208216..3208596)
                     /note="chaperone protein DnaK"
                     /gene="BMETHMGA302782c"
     CDS             complement(3208693..3210048)
                     /note="chaperone dnaK"
                     /gene="BMETHMGA302783c"
     CDS             complement(3210083..3210709)
                     /note="GrpE protein HSP-70 cofactor"
                     /gene="BMETHMGA302784c"
     CDS             complement(3210861..3211757)
                     /note="heat-inducible transcription repressor HrcA"
                     /gene="BMETHMGA302785c"
     CDS             complement(3211979..3213127)
                     /note="oxygen-independent coproporphyrinogen III
          oxidase"
                     /gene="BMETHMGA302786c"
     CDS             complement(3213207..3215042)
                     /note="GTP-binding protein LepA"
                     /gene="BMETHMGA302787c"
     CDS             complement(3215135..3216595)
                     /note="2-oxoglutarate dehydrogenase E1 component"
                     /gene="BMETHMGA302788c"
     fasta_record    3215214..3229332
```

```
                  /note="ConsensusfromContig96 Average coverage:
         466.27"
                  /label=ConsensusfromContig96 Average coverage: 466.27
                  /colour=11
CDS               complement(3216915..3217214)
                  /note="hypothetical protein"
                  /gene="BMETHMGA302789c"
CDS               complement(3217397..3218560)
                  /note="short-chain-specific acyl-CoA dehydrogenase"
                  /gene="BMETHMGA302790c"
CDS               3219497..3220210
                  /note="hypothetical protein"
                  /gene="BMETHMGA302791"
CDS               complement(3220389..3220727)
                  /note="hypothetical protein"
                  /gene="BMETHMGA302792c"
CDS               complement(3220760..3221386)
                  /note="transcriptional regulator, TetR family"
                  /gene="BMETHMGA302793c"
CDS               complement(3221420..3222130)
                  /note="metallo-beta-lactamase family protein"
                  /gene="BMETHMGA302794c"
CDS               3222340..3222666
                  /note="transcriptional regulator SinR"
                  /gene="BMETHMGA302795"
CDS               complement(3223526..3224122)
                  /note="camelysin"
                  /gene="BMETHMGA302796c"
CDS               complement(3224206..3224772)
                  /note="Signal peptidase SipW; Serine peptidase"
                  /gene="BMETHMGA302797c"
CDS               complement(3224812..3225615)
                  /note="hypothetical protein"
                  /gene="BMETHMGA302798c"
CDS               complement(3226085..3226558)
                  /note="hypothetical protein"
                  /gene="BMETHMGA302799c"
CDS               complement(3226628..3227119)
                  /note="methylated-DNA/protein-cysteine
         methyltransferase"
                  /gene="BMETHMGA302800c"
CDS               complement(3227210..3227551)
                  /note="Methylphosphotriester-DNA alkyltransferase"
                  /gene="BMETHMGA302801c"
CDS               3228423..3228998
                  /note="GNAT family acetyltransferase"
                  /gene="BMETHMGA302802"
fasta_record      3229333..3230914
                  /note="ConsensusfromContig97 Average coverage:
         6,400.30"
                  /label=ConsensusfromContig97 Average coverage:
         6,400.30
                  /colour=10
```

```
CDS             complement(3230905..3231150)
                /note=" UTP--glucose-1-phosphate uridylyltransferase"
                /gene="BMETHMGA302803c"
fasta_record    3230915..3234913
                /note="ConsensusfromContig98 Average coverage:
       537.37"
                /label=ConsensusfromContig98 Average coverage: 537.37
                /colour=11
CDS             3231304..3232704
                /note="UDP-glucose 6-dehydrogenase"
                /gene="BMETHMGA302804"
CDS             3232701..3233660
                /note=" NAD-dependent epimerase/dehydratase"
                /gene="BMETHMGA302805"
CDS             3233657..3234799
                /note="spore coat protein"
                /gene="BMETHMGA302806"
fasta_record    3234914..3248577
                /note="ConsensusfromContig99 Average coverage:
       567.42"
                /label=ConsensusfromContig99 Average coverage: 567.42
                /colour=10
CDS             complement(3235449..3236048)
                /note="putative metal-dependent hydrolase"
                /gene="BMETHMGA302807c"
CDS             complement(3236691..3238883)
                /note="YhgE/Pip C-terminal domain protein"
                /gene="BMETHMGA302808c"
CDS             complement(3238957..3239523)
                /note="regulatory protein TetR"
                /gene="BMETHMGA302809c"
CDS             complement(3239749..3241224)
                /note="protoporphyrinogen oxidase"
                /gene="BMETHMGA302810c"
CDS             complement(3241221..3242156)
                /note="ferrochelatase"
                /gene="BMETHMGA302811c"
CDS             complement(3242238..3243284)
                /note="uroporphyrinogen decarboxylase"
                /gene="BMETHMGA302812c"
CDS             complement(3243588..3244280)
                /note="LrgB family protein"
                /gene="BMETHMGA302813c"
CDS             complement(3244280..3244534)
                /note="LrgA family protein"
                /gene="BMETHMGA302814c"
CDS             3245022..3245456
                /note="hypothetical protein"
                /gene="BMETHMGA302815"
CDS             3245484..3245972
                /note="transporter"
                /gene="BMETHMGA302816"
CDS             3246475..3247686
```

```
                /note="amidohydrolase"
                /gene="BMETHMGA302817"
CDS             complement(3247702..3248424)
                /note="hypothetical protein"
                /gene="BMETHMGA302818c"
fasta_record    3248578..3254567
                /note="ConsensusfromContig100 Average coverage:
       577.76"
                /label=ConsensusfromContig100 Average coverage:
       577.76
                /colour=11
CDS             complement(3248846..3252238)
                /note="Glycosyl transferase and polysaccharide
       deacetylase
                fusion protein"
                /gene="BMETHMGA302819c"
CDS             3253937..3254353
                /note="hypothetical protein"
                /gene="BMETHMGA302820"
fasta_record    3254568..3256056
                /note="ConsensusfromContig101 Average coverage:
       9,086.27"
                /label=ConsensusfromContig101 Average coverage:
       9,086.27
                /colour=10
CDS             3254948..3256234
                /note="RNA-directed DNA polymerase"
                /gene="BMETHMGA302821"
fasta_record    3256057..3256924
                /note="ConsensusfromContig102 Average coverage:
       543.77"
                /label=ConsensusfromContig102 Average coverage:
       543.77
                /colour=11
fasta_record    3256925..3258087
                /note="ConsensusfromContig103 Average coverage:
       760.17"
                /label=ConsensusfromContig103 Average coverage:
       760.17
                /colour=10
fasta_record    3258088..3258702
                /note="ConsensusfromContig104 Average coverage:
       570.20"
                /label=ConsensusfromContig104 Average coverage:
       570.20
                /colour=11
CDS             complement(3258683..3259123)
                /note="Spore germination protein, GerC"
                /gene="BMETHMGA302822c"
fasta_record    3258703..3259088
                /note="ConsensusfromContig105 Average coverage:
       1,135.29"
```

```
                     /label=ConsensusfromContig105 Average coverage:
       1,135.29
                     /colour=10
     CDS             3259054..3259971
                     /note="Transporter, drug/metabolite exporter"
                     /gene="BMETHMGA302823"
     fasta_record    3259089..3260169
                     /note="ConsensusfromContig106 Average coverage:
       514.00"
                     /label=ConsensusfromContig106 Average coverage:
       514.00
                     /colour=11
     CDS             3260154..3260684
                     /note="UTP--glucose-1-phosphate uridylyltransferase"
                     /gene="BMETHMGA302824"
     fasta_record    3260170..3272485
                     /note="ConsensusfromContig107 Average coverage:
       539.94"
                     /label=ConsensusfromContig107 Average coverage:
       539.94
                     /colour=10
     CDS             3261063..3262454
                     /note="catalase"
                     /gene="BMETHMGA302825"
     CDS             complement(3262520..3262906)
                     /note="phosphocarrier, HPr family"
                     /gene="BMETHMGA302826c"
     CDS             complement(3263180..3263806)
                     /note="dihydroxyacetone kinase subunit L"
                     /gene="BMETHMGA302827c"
     CDS             complement(3263825..3264820)
                     /note="dihydroxyacetone kinase subunit K"
                     /gene="BMETHMGA302828c"
     CDS             complement(3265073..3266563)
                     /note="glycerol kinase"
                     /gene="BMETHMGA302829c"
     CDS             complement(3266756..3268021)
                     /note="hypothetical protein"
                     /gene="BMETHMGA302830c"
     CDS             complement(3268521..3269720)
                     /note="methionine gamma-lyase"
                     /gene="BMETHMGA302831c"
     CDS             complement(3269924..3270721)
                     /note="hypothetical protein"
                     /gene="BMETHMGA302832c"
     CDS             3271135..3271596
                     /note="Transcriptional regulator MarR family"
                     /gene="BMETHMGA302833"
     CDS             3271785..3272315
                     /note="intracellular protease, PfpI family"
                     /gene="BMETHMGA302834"
     fasta_record    3272486..3281335
```

```
                     /note="ConsensusfromContig108 Average coverage:
       678.93"
                     /label=ConsensusfromContig108 Average coverage:
       678.93
                     /colour=11
CDS                  3273655..3276855
                     /note="CoA-substrate-specific enzyme activase"
                     /gene="BMETHMGA302835"
CDS                  3277090..3278532
                     /note="putative activator of 2-hydroxyglutaryl-CoA
                     dehydratase"
                     /gene="BMETHMGA302836"
CDS                  complement(3279375..3280094)
                     /note="GntR family transcriptional regulator"
                     /gene="BMETHMGA302837c"
CDS                  complement(3280448..3280870)
                     /note=" hypothetical protein"
                     /gene="BMETHMGA302838c"
fasta_record         3281336..3303485
                     /note="ConsensusfromContig109 Average coverage:
       494.91"
                     /label=ConsensusfromContig109 Average coverage:
       494.91
                     /colour=10
CDS                  complement(3281676..3282623)
                     /note="NAD-dependent epimerase"
                     /gene="BMETHMGA302839c"
CDS                  complement(3282709..3283794)
                     /note="8-amino-7-oxononanoate synthase"
                     /gene="BMETHMGA302840c"
CDS                  3284324..3284914
                     /note="hypothetical protein"
                     /gene="BMETHMGA302842"
CDS                  complement(3285031..3285405)
                     /note="hypothetical protein"
                     /gene="BMETHMGA302843c"
CDS                  complement(3285405..3286595)
                     /note="hypothetical protein"
                     /gene="BMETHMGA302844c"
CDS                  3286794..3287597
                     /note="hypothetical protein"
                     /gene="BMETHMGA302845"
CDS                  3288076..3288891
                     /note="hypothetical protein"
                     /gene="BMETHMGA302846"
CDS                  3289095..3290993
                     /note="hypothetical protein"
                     /gene="BMETHMGA302847"
CDS                  complement(3291090..3291887)
                     /note="hypothetical protein"
                     /gene="BMETHMGA302848c"
CDS                  3292179..3293060
                     /note="glycosyl transferase, group 2 family protein"
```

```
                /gene="BMETHMGA302849"
CDS             3293095..3293964
                /note="glycosyl transferase, group 2 family protein"
                /gene="BMETHMGA302850"
CDS             complement(3294082..3294771)
                /note="hypothetical protein"
                /gene="BMETHMGA302851c"
CDS             complement(3294813..3296087)
                /note="maturation of the outermost layer of the
       spore"
                /gene="BMETHMGA302852c"
CDS             complement(3296246..3297208)
                /note="GDP-mannose 4,6-dehydratase"
                /gene="BMETHMGA302853c"
CDS             3297933..3299030
                /note="hypothetical protein"
                /gene="BMETHMGA302854"
CDS             complement(3299181..3300266)
                /note="hypothetical protein"
                /gene="BMETHMGA302855c"
CDS             complement(3300896..3301723)
                /note="glycosyltransferase, group 2 family protein"
                /gene="BMETHMGA302856c"
CDS             3302308..3303363
                /note="UDP-glucose:GDP-mannose dehydrogenase"
                /gene="BMETHMGA302857"
fasta_record    3303486..3315750
                /note="ConsensusfromContig110 Average coverage:
       522.28"
                /label=ConsensusfromContig110 Average coverage:
       522.28
                /colour=11
CDS             complement(3303988..3305361)
                /note="integral membrane transport protein"
                /gene="BMETHMGA302858c"
CDS             3306620..3309049
                /note="Copper-importing ATPase"
                /gene="BMETHMGA302859"
CDS             3309333..3310814
                /note="iron permease FTR1"
                /gene="BMETHMGA302860"
CDS             3310837..3311298
                /note="peptidase M75, imelysin "
                /gene="BMETHMGA302861"
CDS             3311935..3312279
                /note="hypothetical protein"
                /gene="BMETHMGA302863"
CDS             3312473..3313624
                /note="Na(+)/H(+)-K(+) antiporter"
                /gene="BMETHMGA302864"
CDS             3314607..3314978
                /note="hypothetical protein"
                /gene="BMETHMGA302865"
```

```
CDS             3314989..3315372
                /note="hypothetical protein"
                /gene="BMETHMGA302866"
fasta_record    3315751..3316548
                /note="ConsensusfromContig111 Average coverage:
       682.68"
                /label=ConsensusfromContig111 Average coverage:
       682.68
                /colour=10
CDS             complement(3316521..3317207)
                /note="uroporphyrin-III C-methyltransferase"
                /gene="BMETHMGA302867c"
fasta_record    3316549..3324018
                /note="ConsensusfromContig112 Average coverage:
       595.93"
                /label=ConsensusfromContig112 Average coverage:
       595.93
                /colour=11
CDS             complement(3317468..3319087)
                /note="Ferredoxin--nitrite reductase"
                /gene="BMETHMGA302868c"
CDS             complement(3319106..3319705)
                /note="adenylylsulfate kinase"
                /gene="BMETHMGA302869c"
CDS             complement(3319726..3320874)
                /note="sulfate adenylyltransferase"
                /gene="BMETHMGA302870c"
CDS             complement(3320937..3321704)
                /note="phosphoadenosine phosphosulfate reductase"
                /gene="BMETHMGA302871c"
CDS             complement(3322261..3322908)
                /note="uroporphyrinogen-III synthase"
                /gene="BMETHMGA302872c"
CDS             3323494..3323871
                /note="hypothetical protein"
                /gene="BMETHMGA302873"
CDS             3324009..3324410
                /note=" xanthine dehydrogenase, molybdopterin binding
                subunit"
                /gene="BMETHMGA302874"
fasta_record    3324019..3325401
                /note="ConsensusfromContig113 Average coverage:
       930.58"
                /label=ConsensusfromContig113 Average coverage:
       930.58
                /colour=10
CDS             3324609..3325298
                /note="alpha/beta fold family hydrolase"
                /gene="BMETHMGA302875"
fasta_record    3325402..3326181
                /note="ConsensusfromContig114 Average coverage:
       1,162.90"
```

```
                /label=ConsensusfromContig114 Average coverage:
       1,162.90
                /colour=11
fasta_record    3326182..3327120
                /note="ConsensusfromContig115 Average coverage:
       530.80"
                /label=ConsensusfromContig115 Average coverage:
       530.80
                /colour=10
fasta_record    3327121..3329466
                /note="ConsensusfromContig116 Average coverage:
       526.22"
                /label=ConsensusfromContig116 Average coverage:
       526.22
                /colour=11
CDS             3328621..3329349
                /note="conserved hypothetical protein"
                /gene="BMETHMGA302876"
fasta_record    3329467..3329851
                /note="ConsensusfromContig117 Average coverage:
       1,282.00"
                /label=ConsensusfromContig117 Average coverage:
       1,282.00
                /colour=10
CDS             3329496..3329851
                /note="transposase, IS605 OrfB family"
                /gene="BMETHMGA302877"
```

APPENDIX C

Applied and Environmental Microbiology

Genome Sequence of Thermotolerant Bacillus methanolicus: Features and Regulation Related to Methylotrophy and Production of l-Lysine and l-Glutamate from Methanol

Tonje M. B. Heggeset, Anne Krog, Simone Balzer, Alexander Wentzel, Trond E. Ellingsen and Trygve Brautaset

Appl. Environ. Microbiol. 2012, 78(15):5170. DOI: 10.1128/AEM.00703-12.
Published Ahead of Print 18 May 2012.