Independent project in Biology C EX0418, 15HEC
Swedish University of Agricultural Science, SLU
Biotechnology programme, VT-09

Diversity of methanotrophic bacteria in rice fields under different crop rotation regimes

Christine Källman

Supervisor: Nils Högberg
Examinator: Petra Fransson

Sammanfattning

Metanotrofa bakterier använder metan som kol- och energikälla och hittas framför allt i våtmarker som risfält, där produktionen av metan är hög. Metanotrofer oxiderar metan till koldioxid med hjälp av ett enzym som kallas ’perticulate methane monooxygenase’, pMMO. Den aktiva delen av pMMO kodas av subenheten pmoA. Målet med arbetet var att extrahera och isolera den väl bevarade pmoA genen som är en universell biomarkör för metanotrofer. Jordprover togs från ett risfält i sydvästra Vietnam där ris, majs och mungböna odlades. Genen pmoA amplifierades med nested polymerase chain reaction, PCR och ligerades in i en plasmid vektor. Vektorn, med pmoA genen, transformerades in i kompetenta *E.* coli celler för att få ut kolonier med sekvenser av pmoA genen. De kolonier som växte plockades och prover skickades till sekvensering. Resultaten, i form av nukleotidsekvenser, jämfördes genom ett fylogenetiskt träd tillsammans med kända metanotrofer av typ I och II. Det visade sig att proverna innehåll metanotrofer av både typ I och II, som formade två kluster i det fylogenetiska trädet. Det fanns mer metanotrofer i behandlingen med samtliga plantor; ris, mungböna och majs än i behandlingen med endast ris. Minst fanns det i jorden där endast ris odlades. I materialet fanns även nya sekvenser som inte hittats tidigare.

Aim

The aim of this study was to extract and isolate DNA from methanotrophic bacteria originated from rice fields and obtain data that can be compared using a phylogenetic analysis. Another purpose was to link the sequence diversity to different rice field crop rotation management strategies. Furthermore, I wanted to learn more about the biological methods that are regularly used in lab, like polymerase chain reaction, cloning, transformation and sequencing.

Introduction

Background
Methane is produced naturally in nature when methanogens digest plant material in absence of oxygen. Methanotrophs can oxidize the methane that the methanogens produce, which is a process that reduce the amount of methane released to the atmosphere. They can also oxidize the methane that already is in the atmosphere. Methanotrophs are a hot topic for researches since methane is a more potent greenhouse gas than carbon dioxide. Furthermore, methanotrophs are able to detoxify organic toxic components called chlorinated hydrocarbons from contaminated soils. Methanotroph's ability to degrade for example chlorinated hydrocarbon, make them suitible to use as biomediators (Knief et al., 2005, McDonald et al., 2008).

Methanogens
Methanogens were discovered in the 1970s and are common in wetlands but can also be found in the guts of animals and humans. Methanogens are microbes that produce methane as a biproduct from its metabolism. Methane is produced from different substrates like H_2, CO_2 or acetate, when there are no or little oxygen (Conway et al., 2009). Methanogens that uses carbon dioxide as the source of carbon and hydrogen as energy source are called *hydrotropic*. The carbon and hydrogen reacts to reduce the carbon and produce methane The methane is then used in cellular pathways that generate ATP. The ones that use acetate as both the carbon and energy source are called *acetotrophic*. Methanogenes, mostly acetotrophic, are also used in the anaerobic degeneration of organic material to biogas but a problem here is that they are pH- and temperature sensitive (Anna Schnürer 2008, Hellström 2001).

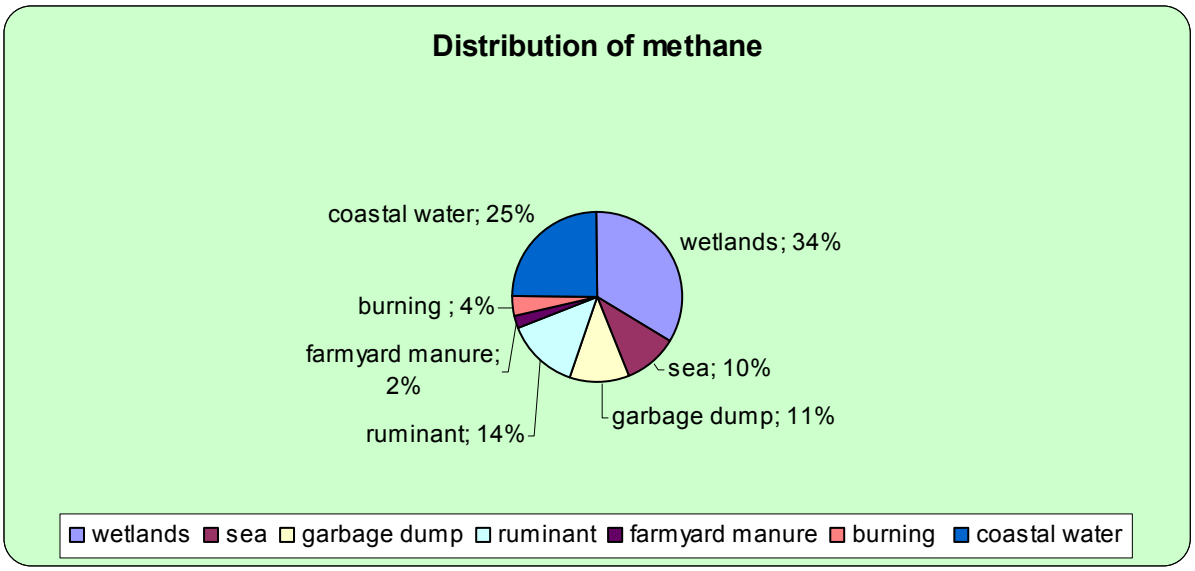


Figure 1. Picture over the distribution of methane sources where wetlands and coastal water are the main sources and stands for almost 60% of the methane production (Rodhe et al., 1995).

Methanotrophic bacterias

Bacteria that are able to metabolize methane as their carbon and energy source are called methanotrophs. Methanotrophs can usually be found in soils with a high production of methane, like swamps, rice paddies and bogs, and does not require oxygen to grow. Wetlands are the biggest source of methane production, which can be seen in *figure 1*. Coastal water stands for approximately 25% of methane production and ruminants for 14%. Burning only stands for 4 % of the production (Rodhe et al., 1995).

Under aerobic conditions, aerobic methanotrophs will oxidize methane. Anaerobic methanotrophs will oxidize methane under anaerobic conditions which is rarer. Aerobic methanotrophic bacteria are divided into two taxonomic groups, type I and II, depending on their cell morphology, metabolism and phylogeny. Type I methanotrophs appears to be favoured in environments with limiting methane, and type II in environment with high levels of methane (Hanson RS et al., 1996). Type I methanotrophs are members of the class *Gammaproteobacteria* and type II methanotrophs belongs to the class *Alphaproteobacteria*. Type I include the genera *Methylobacter*, *Methylomicrobium*, *Methylomonas*, *Methylocaldum*, *Methylosphaera*, *Methylothermus*, *Methylosarcina*, *Methylohalobius*, *Methylosoma*, and *Methylococcu*s. Type II methanotrophs include the genera *Methylocystis*, *Methylosinus*, *Methylocella*, and *Methylocapsa* (McDonald et al., 2008).

Particulate methane monooxygenase

There are several steps to oxidize methane to carbon dioxide, which can be seen in *figure 2*. The first step is to convert methane to methanol which is done by the enzyme methane monooxygenase, a key enzyme in methane oxidation. This enzyme exists in two forms, a soluble cytoplasmic form, sMMO, and a particulate membrane bound form, pMMO. The pMMO is the most common form and is present in all methanotrophs except in one genus called *Methylocella*. The sMMO form is not so common and exists only in some strains of methanotrophs (McDonald et al., 2008). The enzyme pMMO contains both iron and copper. The enzyme is encoded by three genes which is responsible for three integral membrane polypeptides that build up the enzyme. The genes are called pmoA, pmoB and pmoC. pmoA

is a 27 kDa subunit of pMMO and encodes the active site polypeptide of pMMO. It has been shown that pMOA is a highly conserved gene among methanotrophs, both in type I and II (Tchawa et al., 2003).

Since almost all methanotrophs have the enzyme pMMO, they also have the highly conserved gene pmoA that can be used as a universal marker gene for methanotrophs (Knief et al., 2005). Parallels can be drawn between the phylogeny of pmoA and 16s RNA for pMMO since no horizontal transfer of pmoA is yet discovered among methanotrophs (Horz et al., 2005). Horizontal gene transfer is when an organism incorporates DNA from another organism without being the offspring. The pmoA gene is now used in studies of methane oxidation activities and the methanotrophic community composition in different types of land like cornfields and forest (Knief et al., 2005). The pmoA gene can also be used to see what type of methanotrophic bacteria is in the soil and how it is related to another methanotroph and if they are involved in atmospheric methane consumption (Horz et al., 2004).

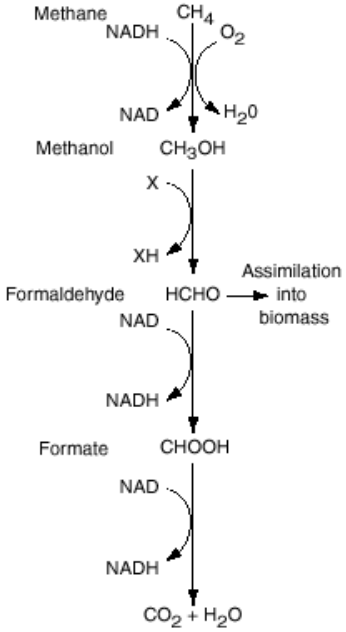


Figure 2. The metabolic pathway of methanotrophic bacteria, showing how the methane is oxidized to carbon dioxide (Brigmon, 2001).

Polymerase chain reaction, PCR

PCR is a biological technique used for copying DNA sequences by using a heat stable DNA polymerase followed by temperature changes. Usually the PCR is run for around 30 repeated cycles. The first step in PCR is called *initial denaturation step* and here the reaction is heated to around 95°C to reduce non-specific amplification. The next step is called the *denaturation step* where the reaction is held at around 95°C. The hydrogen bonds between complementary bases in the double stranded DNA will break and form two single stranded DNA sequences. After denaturation there is an *annealing step* where the temperature is decreased to around 60°C and the primers are attached to the DNA template. The primers are short DNA fragments that are complementary to the DNA sequence of interest. When the primers attach to the template, hydrogen bonds will be formed. The DNA polymerase will recognize the primer and bind to the template and begin DNA synthesis.

The next step is called *elongation step* and is performed in temperatures just over 70°C, where the DNA polymerase has its optimum activity. The DNA polymerase will add a dNTP to the new strand, which are complementary to the template's nucleotides. Under optimum conditions in the extension step, the DNA polymerase can double the amount of DNA. The last step is called *final elongation*, where the temperature is held at around 70°C after the last PCR cycle to be sure that the single stranded DNA template has a complementary newly synthesized DNA strand. The reaction is then holding for around 5°C that will function as short time storage of the CPR product.

Nested PCR is a variant of the regular PCR reaction described above. Nested PCR is used to eliminate contaminations of non-specifically amplified PCR products and ensure that the

primers have bound to the right region of the DNA template by using two sets of primers. The first pair of primers will amplify a DNA fragment in the first run of PCR. This PCR product is then used as the DNA template in the next PCR run with the second pair of primers. This second pair of primers is nested or binds inside the first PCR product which makes the second PCR product shorter (contains less basepairs) then the first product. So if there were unspecific binding of the first pair of primers that leads to amplification of a wrong DNA fragment, the second pair of primers will not bind at all, leading to no PCR product. Nested PCR is used for very low DNA koncentration and if there are problems with the primers (www.pcrstation.com/nested.pcr).

Touch down PCR is an improved PCR method which will reduce the probability for the primers to amplify nonspecific sequences. The specificity of a PCR reaction is determinded by the temperature when primers anneal to the template. When annealing temperature is just below the melting temperature for the primers, only very specific paring occurs between the primers and the template. The specificity of the primers will decrease with lower annealing temperatures. The program uses varying annealing temperatures in one cycle and begins with a high temperature, just below the melting temperature of the primers. The anneling temperature is then decreasing with the increasing number of cycles. Primers anneal to the template at the highest temperature they can tolerate, which makes the first sequence amplified, most likely the sequence of interest. This sequence will then be amplified exponentially at lower annealing temperatures (Don et al., 1991).

Phylogenetic trees

Comparing morphological characters is the classical way to investigate relationships between species. Species can also be compared on their molecular data, more specific the nucleotide sequence, and that is what have been done in this study. The evolutionary relationship among genes can be described with a a dendogram which is phylogenetic tree (Salemi et al., 2003). Phylogenetic trees can be constructed based on obtained data and the purpose of the analysis. A generic term that describes the existing data, the so called external nodes, which can be a family of organisms or a set of related genes are often called operational taxonomic units, OTUs. The internal nodes, the hypothetical progenitors of OTUs, are called hypothetical taxonomic units, HTUs. A cluster is a group of taxa that have a common ancestor. A cladogram is a phylogenetic tree where the length of the horizontal branches function as time axes. A phylogram is a phylogenetic tree where the length of the branches function is based on the number of evolutionary changes.

Bootstrap analysis can be used when making phylogenetic trees to estimate statistical errors. This resampling method samples columns from the original alignment, by randomly choosing columns from it, and analyses if the samples data gives the same tree. One column can be selected more than once or not at all until a new set of sequences are obtained with the same length as original. For every reproduced dataset, a tree is constructed, and the proportion of each clade among the the datasets is computed. These values are shown as bootstrap values on the tree (Salemi et al., 2003).

Material and methods

64 soil samples was taken in Mekong Delta region, in southwestern Vietnam. Under intensive rice cultivation rice can be harvested three times per year but this management system is not

sustainable so here crop rotation is tested as an alternative according to the following schedule with three crops per year:
R-R-R means treatment with rice-rice-rice
R-Mg-R means treatment with rice-mungbean-rice
R-M-R means treatment with rice-maize-rice
R-Mg-M means treatment with rice-mungbean-maize

All samples were grown with rice when samples were collected.

R-M-R 57-64 II_4	R-R-R 25-32 I_4	R-M-R 49-56 II_3	R-Mg-R 89-96 III_4
64 63 62 61 60 59 58 57	32 31 30 29 28 27 26 25	56 55 54 53 52 51 50 49	96 95 94 93 92 91 90 89
R-Mg-R 81-88 III_3	R-M-R 41-48 II_2	R-Mg-M 121-128 IV_4	R-R-R 17-24 I_3
88 87 86 85 84 83 82 81	48 47 46 45 44 43 42 41	128 129 126 125 124 123 122 121	24 23 22 21 20 19 18 17
R-R-R 9-16 I_2	R-Mg-M 113-120 IV_3	R-Mg-R 73-80 III_2	R-Mg-M105-112 IV_2
16 15 14 13 12 11 10 9	120 119 118 117 116 115 114 113	80 79 78 77 76 75 74 73	112 111 110 109 108 107 106 105
R-Mg-M 97-104 IV_1	R-Mg-R 65-72 III_1	R-R-R 1-8 I_1	R-M-R 33-40 II_1
104 103 102 101 100 99 98 97	72 71 70 69 68 67 66 65	8 7 6 5 4 3 2 1	40 39 38 37 36 35 34 33

DNA extraction from soil samples
DNA extraction was done in two ways with 12 soil samples to see which extraction method was most effective and gave best results. 0.5g soil was taken from each sample for extraction. The first DNA extraction method was done with a kit called FastDNA® SPIN Kit for Soil (Q-BIO gene), according to the instructions in the manufacturer's manual. The second extraction method was made according to the method in Griffith et al., 2000. Both extraction methods are designed to extraxt PCR-ready genomic DNA.

There are three main steps in the DNA extraction. First step is lysing of the cells, which means that the membrane of the bacteria will be broken open. Then buffers and other reagents are added to homogenize the sample with a little RNA contamination as possible. The last step is DNA purification and elution.

PCR amplification of extracted DNA

The DNA extract was amplified with nested PCR, which means two separate PCR reactions with different pair of primers. The primers where ordered from TAG Copenhagen A/S and designed to amplify the pmoA gene:

Primers
A189F: 5'GGN GAC TGG GAC TTC TGG 3' (forward, 18 bases)
A682R: 5'GAA SGC NGA GAA GAA SGC 3' (reverse, 18 bases)
mb661: 5' CCG GMG CAA CGT CYT TAC C 3' (reverse, 19 bases)

A189F/A682R creates a product size of 525bp
A189F/mb661 creates a product size of 510bp (McDonald IR et al., 2008)

Codes of the International Union of Biochemistry (IUB codes)
N = A+C+G+T
S= C+G
Y= C+T

The first PCR reaction contained following components:
Dream Taq polymerase
10x buffer
dNTP
Forward primer: A189F
Reverse primer: A682R
Template (DNA extraction)
+ De-ionized water to a total volume of 10µl / reaction

The second PCR reaction contained following components:
Dream Taq polymerase
10x buffer
dNTP
Forward primer: A189F
Reverse primer: mb661
Template (PCR product from the first PCR run)
+ De-ionized water to a total volume of 20µl / reaction

A touch down PCR program was used similar as in the method described in Claudia K et al., (2005), with annealing temperatures 62ºC to 52ºC.

"Touch down PCR"
94ºC for 2min (*initial denaturation step)*
94ºC for 0,20min (*denaturation step)*
62ºC for 0,30min (*annealing step)*
72ºC for 2min (*elongation step)*
94ºC for 0,20min (*denaturation step)*
52ºC for 0,30min (*annealing step)*
72ºC for 2min (*elongation step)*
4ºC for 10min (*storage)*

Verification of DNA with gel electrophoresis

After the DNA have been amplified with PCR, the content was verified with gel electrophoresis. Gel electrophoresis is a method that separates DNA according to their size. An electric current pulls the negative charged DNA to a anod and smaller DNA fragment will migrate faster through gel than larger fragment.

The agarose gel was prepared with 12% agarose (1.8g agarose/ 150 ml buffer). The gel was run with 400mA and 150 volt for approximately one hour.

Cloning and transformation with competent cells

The PCR product from the 64 samples where assembled into groups of four, (where the four samples where taken from the same square in the field):

Table 1. PCR samples assembled into tubes before cloning.

Tube	Samples
1	19, 20, 21, 22
2	35, 36, 37, 38
3	59, 60, 61, 62
4	3, 4, 5, 6
5	11, 12, 13, 14
6	27, 28, 29, 30
7	43, 44, 45, 46
8	51, 52, 53, 54
9	67, 68, 69, 70
10	75, 76, 77, 78
11	83, 84, 85, 86
12	91, 92, 93, 94
13	99, 100, 101, 102
14	107, 108, 109, 110
15	115, 116, 117, 118
16	123, 124, 125, 126

For cloning and transformation, a kit was used, called TOPO TA Cloning® from Invitrogen, according to the manufacturer's instructions. The plasmid vector used for cloning was pCR®2.1-TOPO®, and an overview can be seen in *figure 3*. The 16 samples of PCR products where first incubated with the vector. Since *Taq* polymerase has a nontemplate-dependent terminal transferase activity, it adds an "A" (a single deoxyadenosine) to the 3´ ends of the PCR product. The plasmid vector, supplied with the kit, has an overhang of "T" (a single deoxythymidine) at the 3´end which makes an easy ligation of the PCR product into the vector. Topoisomerase I is covalently bound to the vector and activates the vector and then be released after PCR product has been ligated.

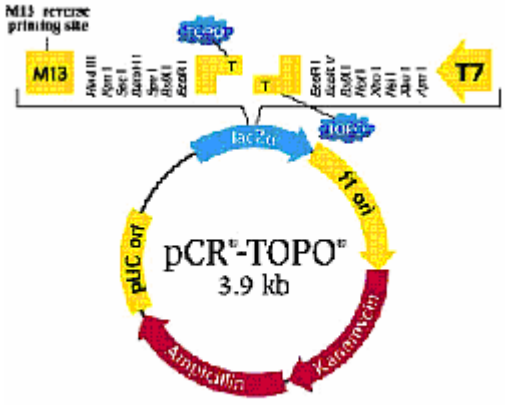


Figure 3 Picture showing the plasmid vector pCR®2.1-TOPO®, which was used to clone PCR products (www.invitrogen.com).

The plasmid vectors with PCR products are then transformed into chemically competent E. *coli* cells, supplied with the kit. The transformation is done because the PCR oproducts,

ligated into the vectors, consist of a mix of sequences. Ampicillin was used as a resistance marker and the transformants were grown on LB-plates containing ampicillin. After incubation, white colonies could be selected for sequencing, which contain only one type of sequence per colony. Transformants were then visualized by gel electrophoresis with primer pair M13:
M13 forward primer: 5´G TAA AAC GAC GGC CAG 3´
M13 reverse primer: 5´CAG GAA ACA GCT ATG AC 3´
The colonies were then prepared for sequencing. A kit was used called ABI PRISM® BigDye™ Terminator Cycle Sequencing Ready Reaction Kit from Applied Biosystems, according to the manufacturer's instructions ("Precipitating Microcentrifuge Tubes").
The samples where send to the sequencing company "Macrogen", for Sanger cycle sequencing.

Data analysis
Sequence quality was manually checked and the closest matching sequences in the NCBI database to my sequences were identified via the BLAST tool and introduced into the dataset. All sequences were aligned in the programs Clustal-W and MUSCLE. The alignment from the programme MUSCLE was used after a manual comparison of the results. The aligned dataset was formatted in PHYLIP-format and a Maximum likelihood approach was used to construct a phylogenetic tree. The branches were evaluated with bootstrap replicates.

Results

Results from DNA extraction

The two different extraction methods where tested with the same 12 samples, and the results can be seen in *figure 3*. The first upper row shows PCR product from the first PCR reaction in nested PCR (with primers A189F and A682R) and the second row shows PCR products from second PCR reaction (with primers A189F and mb661). The number and orientation of the samples are the same in both rows. The wells, counted from the left contained; 100bp DNA ladder the next 12 wells contained sample 19, 20, 21, 22, 35, 36, 37, 38, 59, 60, 61 and 62. These wells are DNA extracted with the FastDNA® SPIN Kit for Soil, from Q-BIO gene. The following 12 wells contained the same samples in the same orientation and were extracted according to the method in Griffith et al., 2000. The next well contained a negative control with water, and the last well contained 100bp DNA ladder. The gel electrophoresis picture shows that FastDNA® SPIN Kit for Soil, worked for all samples, and the PCR product was around 500bp, while the other method only worked for two samples. So the kit was then used for all the following samples.

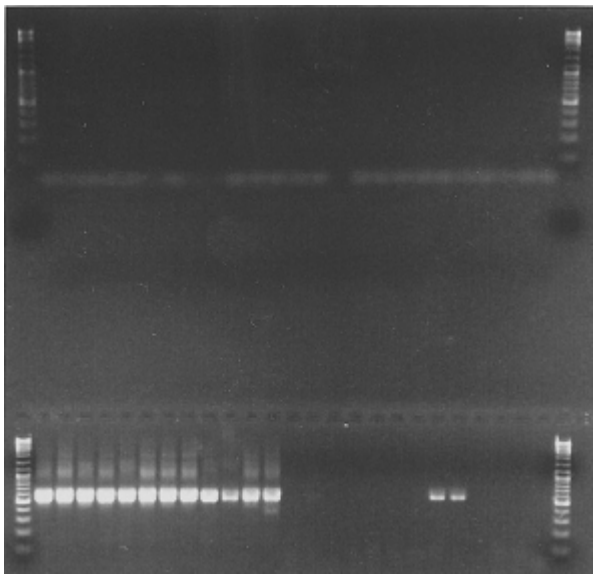


Figure 3. Gel electrophoresis photo showing PCR products around 500bp, after DNA extraction from 12 soil samples with two different methods. The wells in upper rows, counted from the left contained; 100bp DNA ladder the next 12 wells contained sample 19, 20, 21, 22, 35, 36, 37, 38, 59, 60, 61 and 62. These samples were extracted with FastDNA® SPIN Kit for Soil, from Q-BIO gene. The following 12 wells contained the same samples and were extracted according to the method in Griffith et al., 2000. The upper row hows PCR products with primers A189F and A682R. The next well was a negative control with water, and the last well contained 100bp DNA ladder. The lower row contains the same samples, in the same orientation, as in the upper row but the the DNA were amplified with primers A189F and mb661. The photo shows that the kit was most effective for DNA extraction.

After the 12 first samples had been extracted, the following 52 samples were extracted with the same kit. A gel electrophoresis of the PCR products, amplified with primers A189F and mb661, can be seen in *figure 4.* The wells counted from upper row contained: 100bp DNA ladder, sample 3, 4, 5, 6, 11, 12, 13, 14, 27, 28, 29, 30, 43, 44, 45, 46, 51, 52, 53, 54, 67, 68, 69, 70, 75, 76, 77, 78 and 100bp DNA ladder. The downer row counted from left contained, a 100bp DNA ladder, sample 83, 84, 85, 86, 91, 92, 93, 94, 99, 100, 101, 102, 107, 108, 109, 110, 115, 116, 117, 118, 123, 124, 125, 126 and the second lowest well contained a negative control with water and the last well a 100bp ladder. 47 of 52 samples were amplified. The picture shows that the DNA extraction had worked well for almost all samples and contained a PCR product of around 500bp. Although, a thin or no band can be seen in sample 3, 5, 45, 51, 68 and 86 that indicates that less o no DNA had been obtained from the extraction. The PCR products were put in the freezer for later use of cloning and transformation.

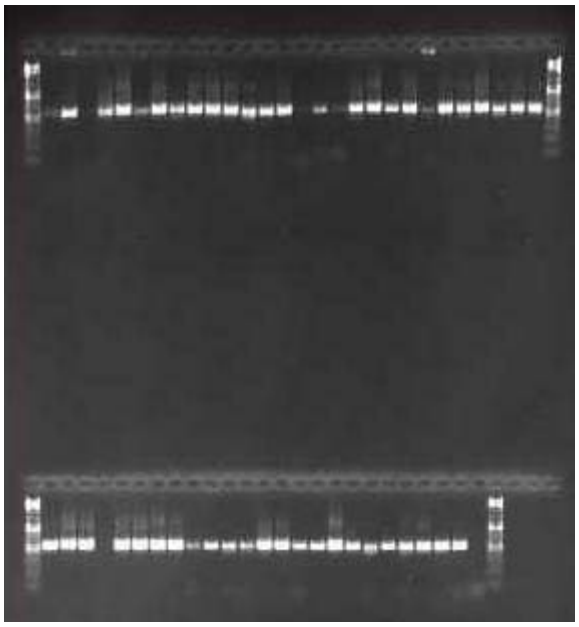


Figure 4. Gel electrophoresis picture showing PCR products of DNA extracted with FastDNA® SPIN Kit for Soil, from Q-BIO gene, with primers A189F and mb661. The wells in upper rows, counted from the left contained; 100bp DNA ladder, sample 3, 4, 5, 6, 11, 12, 13, 14, 27, 28, 29, 30, 43, 44, 45, 46, 51, 52, 53, 54, 67, 68, 69, 70, 75, 76, 77, 78 and 100bp DNA ladder. The wells in the downer row, counted from the left contained; 100bp DNA ladder, sample 83, 84, 85, 86, 91, 92, 93, 94, 99, 100, 101, 102, 107, 108, 109, 110, 115, 116, 117, 118, 123, 124, 125, 126 and the second lowest well contained a negative control with water and the last well a 100bp DNA ladder. Overall, the picture shows a god DNA extraction of the sample, that contained a PCR product of around 500bp.

Results from cloning and transformation

The first cloning and transformation did not work, transformed colonies appeared only on one plate. Therefore, the PCR reaction followed by cloning and transformation was made again with control samples, supplied with the kit. A gel electrophoresis picture of the control samples can be seen in *figure 5*. The wells, counted from left contained: 100bp DNA ladder, control sample 1, 2, 3 a negative control with water and a 100bp DNA ladder. The control samples all contained the same DNA. The picture shows that the PCR reaction had worked well and the PCR product has a lenght of around 500bp and therefore they were immediately cloned and put in the freezer for transformation.

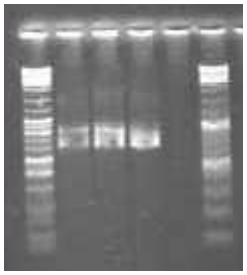


Figure 5. A gel electrophoresis photo of PCR product from a control, supplied with the kit. The wells, counted from left contained: 100bp DNA ladder, control sample 1, 2, 3 a negative control with water and a 100bp DNA ladder. The PCR product are around 500bp.

The PCR reaction with primers A189F and mb661 was made again for all samples. This was to see if the PCR products had to be resh for the cloning. A gel electrophoresis picture of the

PCR products can be seen in *figure 6*. The first row, counted from left contained; 100bp DNA ladder, sample 19, 20, 21, 22, 35, 36, 37, 38, 59, 60, 61, 62, 3, 4, 5, 6, 11, 12, 13, 14, 27, 28, 29, 30, 43, 44, 45, 46 and 100bp DNA ladder. The second row counted from left, contained; 100bp DNA ladder, sample 51, 52, 53, 54, 67, 68, 69, 70, 75, 76, 77, 78, 83, 84, 85, 86, 91, 92 and 100bp DNA ladder. The third row counted from the left, contained; 100 bp DNA ladder, sample 93, 94, 99, 100, 101, 102, 107, 108, 109, 110, 115, 116, 117, 118, 123, 124, 125, 126 and a negative control with water. *Figure 6* shows that almost all samples contained DNA but there were less in sample 45 and 86, verified by the thin band in the gel. The PCR products were around 500bp and immediately cloned and the frozen before transformation.

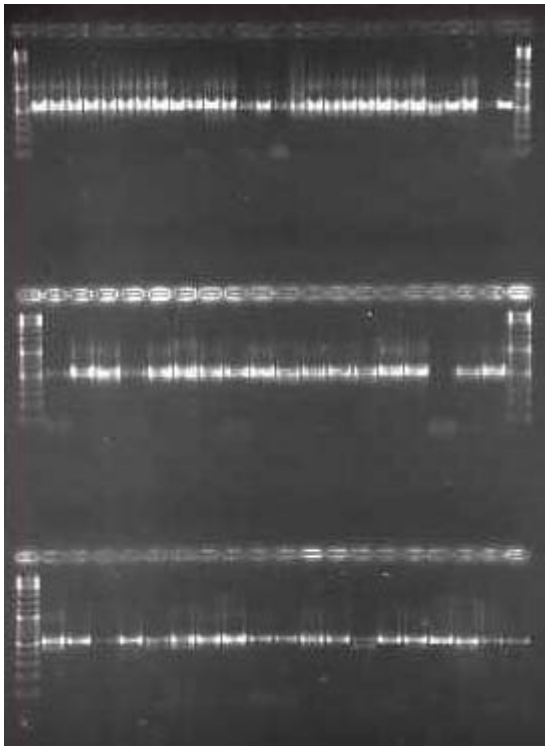


Figure 6. A gel electrophoresis photo showing PCR products from all samples, amplified with primers A189F and mb661. The first row, counted from left contained; 100bp DNA ladder, sample 19, 20, 21, 22, 35, 36, 37, 38, 59, 60, 61, 62, 3, 4, 5, 6, 11, 12, 13, 14, 27, 28, 29, 30, 43, 44, 45, 46 and 100bp DNA ladder. The second row counted from left, contained; 100bp DNA ladder, sample 51, 52, 53, 54, 67, 68, 69, 70, 75, 76, 77, 78, 83, 84, 85, 86, 91, 92 and 100bp DNA ladder. The third row counted from the left, contained; 100 bp DNA ladder, sample 93, 94, 99, 100, 101, 102, 107, 108, 109, 110, 115, 116, 117, 118, 123, 124, 125, 126 and a negative control with water. The samples had a PCR product around 500bp.

After cloning and transformation into competent *E*. coli cells, there where colonies on plate 2, 3, 4, 5, 6, 10, 11, 12, 13, 14, 15 and 16. The original 64 samples were, as described earlier divided into groups of 4. This can be seen in *table* 1 in "Material and Methods". There were only few colonies on each plate. The pmoA gene was then amplified in PCR with primer pair M13F and M13R. A gel electrophoresis picture can be seen in *figure 7.* The wells counted from left contained 100bp DNA ladder, colonies from plate 2, 3, 4, 5, 6,10, 11, 12, 13, 14, 15, 16 and a 100bp DNA ladder. Unfortunately, more than one colony were put in the same tube, that could be resulting in a mixture of sequences. Also, only a tenth of the total volume in the PCR reaction was template. As can be seen in *figure 7*, no band around 500bp is visible for any sample.

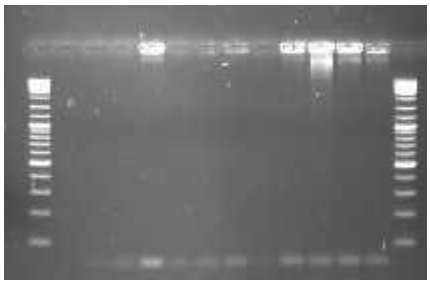


Figure 7. Gel electrophoresis picture showing PCR product of transformants, amplified with primer pair M13F and M13R. The wells counted from left contained 100bp DNA ladder, colonies from plate 2, 3, 4, 5, 6,10, 11, 12, 13, 14, 15, 16 and a 100bp DNA ladder. No visible band can be seen in the gel that correspond to the pmoA gene (around 500bp).

The transformation was made again for all samples. Now the template volume was increased to half of the total volme in the PCR reaction. The transformation went better this time and together with the old transformants,
a total of 60 colonies could be picked from 11 plates. One colony could be picked from plate 2, one from plate 4, one from plate 5, one from plate 6, three from plate 10, three from plate 11, three from plate 12, eleven from plate 13, twelve from plate 14, sixteen from plate 15, eight from plate 16. A gel electrophoresis picture of the PCR products of the transformants can be seen in figure *figure 6*. PCR product were amplified with primer pair M13F and M13R. The wells in the upper row counted from the left, contained; 100bp DNA ladder, sample 5, 10a, 10b, 10c, 11a, 11b, 11c, 12a, 12b, 12c, 13a, 13b, 13c, 13d, 13e, 13f, 13g, 13h, 13i, 13j, 13k, 14a, 14b, 14c, 14d, 14e, 14f, 14g and a 100bp DNA ladder. The second row counted from left, contained; 100bp DNA ladder, sample 14h, 14i, 14j, 14k, 14i, 15a, 15b, 15c, 15d, 15e, 15f, 15g, 15h, 15i, 15j, 15k, 15l, 15m, 15n, 15o, 15p, 16a, 16b, 16c, 16d, 16e, 16f, 16g and a 100bp DNA ladder. The third row counted from left, contained; 100bp DNA ladder, 16h, 2, 4, 6, negative control with water and a 100bp DNA ladder. *Figure 8* shows that 16 samples contained the pmoA gene (band around 500bp). The band are higher than expected, which means that the amplified DNA has a bigger size than the pmoA gene, probably there are some sequences from the plasmid vector that have been amplified as well. In some samples, the plasmid vector had been amplified and therefore there is a smaller band around 200bp. The samples that were send to sequencing were; 4, 10a, 10b, 10c, 12b, 12c, 13a, 13b, 13d, 15b, 15d, 15e, 15i, 16b, 16c and 16e.

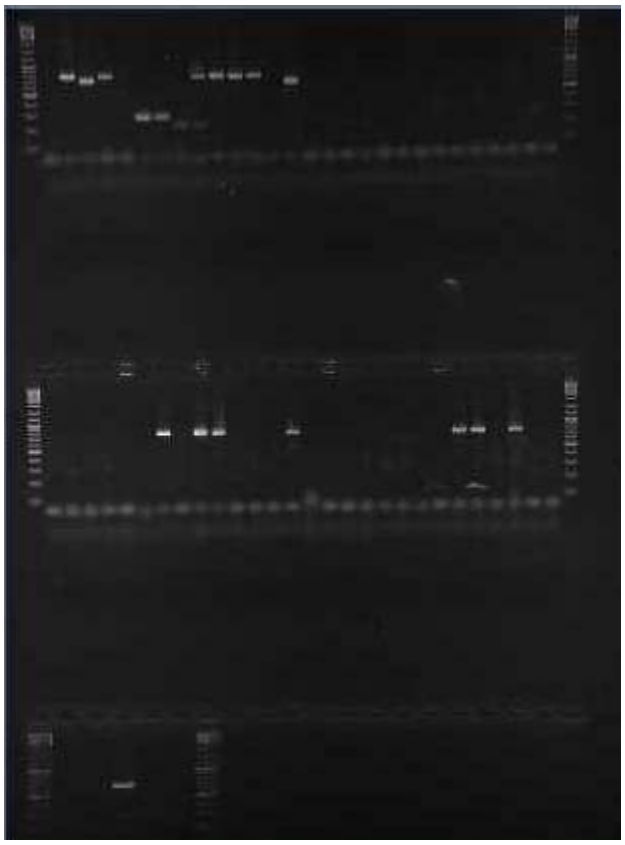


Figure 8. A gel electrophoresis photo of the PCR products of transformants, amplified with primers M13F and M13R. A total of 60 clones were selected from 11 plates. 16 samples contained the pmoA gene around 500bp. The wells in the upper row counted from the left, contained; 100bp DNA ladder, sample 5, 10a, 10b, 10c, 11a, 11b, 11c, 12a, 12b, 12c, 13a, 13b, 13c, 13d, 13e, 13f, 13g, 13h, 13i, 13j, 13k, 14a, 14b, 14c, 14d, 14e, 14f, 14g and a 100bp DNA ladder. The second row counted from left, contained; 100bp DNA ladder, sample 14h, 14i, 14j, 14k, 14i, 15a, 15b, 15c, 15d, 15e, 15f, 15g, 15h, 15i, 15j, 15k, 15l, 15m, 15n, 15o, 15p, 16a, 16b, 16c, 16d, 16e, 16f, 16g and a 100bp DNA ladder. The third row counted from left, contained; 100bp DNA ladder, 16h, 2, 4, 6, negative control with water and a 100bp DNA ladder.

The results from sequencing

All 16 samples sent to sequencing contained the pmoA gene.
(Any sequence that had been verified to belong to the plasmid vector has been deleted in the samples).

Sample 4

GGGGACTGGGACTTCTGGCCTTCTTCTCCGCCTTCCGTTCGGCGCCGTTTTCGCGGCTCTGGGCCTCCTGATTGGCGAGTGGA
TCAACCGCTACGTCAACTTCTGGGGCTGGACCTATTTCCCGATCAGCCTCGTGTTCCCGTCCGCTCTGATCGTTCCGGCGATC
TGGCTTGACGTGATCCTGCTTCTGTCGGGTTCCTATGTGATCACGGCGGTTGTCGGTTCGCTGGGCTGGGGTCTGCTGTTCTA
CCCGAACAACTGGCCGGCGATCGCCGCCTTCCACCAGGCGACCGAGCAGCATGGTCAGCTGATGACCCTGGCTGACCTGATCG
GCCTCCACTACGTCCGCACGTCGATGCCGGAATACATCCGCATGGTCGAGCGCGGCACGCTGCGCACGTTCGGTAAGGACGTT
GCTCCGGAAGGGCGAATTCTGCAGATATCCATCACACTGGCGGCCGCTCGAGCATGCATCTAGAGGGCCCCNTGNANANCAGT
ATAGGACNATNNNN

A DNA sequence trace can be seen in *Appendix 1*.

Sample 10a

GGGGACTGGGACTTCTGGACCGACTGGAAAGACCGTCGTCTGTGGGTAACCGTATTGCCAATCGTTGGTATTACATTCCCGGC
TGCGGTTCAAGCAGTTCTTTGGTACCGTTACCGTCTGCCATTCGGCGCTATGCTGGCTGTATTAGGTCTGCTGTTCGGCGAAT

GGGTCAACAGATACTTCAACTTCTGGGGATGGACTTACTTCCCGGTTAACTTTGTATTCCCATCACAATTCGTTCCAGGCGCA
CTCGTTCTGGACGTGATCCTGATGTTGTCAAACAGCATGCAGTTGACTGCTGTGCTGGGCGGCTTGGCTTATGGTCTGTTGTT
CTATCCTGGCAACTGGCCTGTCATCGCTCCATTGCACGTGCCAGTTGAATACGACGGCATGGTAATGACCCTGGCTGACTTGC
AAGGTTACCACTACGTAAGAACTGGTACTCCAGAGTACATCAGAATGGTTGAAAAAGGTACTTTGAGAACATTCGGTAAAGAC
GTTGCTCCGGAAGGGCGAATTCCAGCACACTGGCGGCCGTTACTAGTGGATCCGAGCTCGGTACCAAGCTTGGCGTAATCATG
GTCATNCCATGTTTTCCNA

A DNA sequence trace can be seen in *Appendix 1*.

Sample 10b

GGGGACTGGGACTTCTGGCCTTCTTCTCCGCGTTCTGTTCGGCGCCGTTTTCGCGGCTCTGGGCCTCCTGATCGGCGAGTGGA
TCAATCGCTACGTCAACTTCTGGGGCTGGACCTATTTCCCGATCAGCCTCGTGTTCCCGTCCGCTCTGATCGTCCCGGCGATC
TGGCTCGACGTGATCCTGCTTCTGTCGGGTTCCTATGTGATCACGGCGGTTGTCGGTTCGCTGGGCTGGGGTCTGCTGTTCTA
CCCGAACAACTGGCCGGCGATTGCGGCGTTCCACCAGGCGACCGAGCAGCATGGTCAGCTGATGACGCTTGCTGACCTGATCG
GCCTCCACTACGTCCGCACGTCGATGCCGGAATACATCCGCATGGTCGAGCGCGGCACGCTGCGCACGTTCGGTAAGGACGTT
GCTCCGGAAGGGCGAATTCCAGCACACTGGCGGCCGTTACTAGTGGATCCGAGCTCGGTACCAAGCTTGGCGTAATCATGGTC
ATGNTTGTTTTCCAAAANNCGGTNCGAACCCTGCTCGCTACCTAGCTCGCCGCAATCGTGTTCTGGTATGTTTCCAAAATTAN
NAATTATGTTTTTA

A DNA sequence trace can be seen in *Appendix 2*.

Sample 10c

GGCGACTGGGACTTCTGGGTTGACTGGAAGGATCGTCGTTTGTGGGTGACGGTGGTGCCGATCGTGCTGGTAACTTTCCCGGC
TGCGGTACAGGCCTTCTTGTGGGAGCGTTTCCGTCTGCCTTGGGGCGCGACGATTTGCGTACTGGGCCTGCTGTTTGGTGAGT
GGGTCAACCGCTACTTCAACTTCTGGGGCTGGACCTACTTCCCGATCACCCTGTGCTTCCCGTCGCAGATCGTACCGGGCGCC
ATCCTGCTCGACACGGCTCTGATGCTGAGCACCAGCTACCTGTTCACCGCGATCGTCGGTGCGATGGCTTGGGGCTTGGTTTT
CTATCCGGGCAACTGGCCGGTGATTGCACCGTACCACGTGCCGGTAGAGTACAACGGCATGCTGATGTCGGTTGCCGACCTGC
TGGGTTACCACTATGTCCGTACGGGTACGCCTGAGTACATCCGTATGGTCGAGAAGGGCACCCTGCGTACCTTCGGTAAAGAC
GTTGCTCCGGAAGGGCGAATTCTGCAGATATCCATCACACTGGCGGCCGCTCGAGCATGCATCTAGAGGGCCCNNNNNNANCA
GCTAGTANNTANNN

A DNA sequence trace can be seen in *Appendix 2*.

Sample 12b

NNNNNTCGNNNATNCTGNTTNNTGGGCCCTCTAGATGCATGCTCGAGCGGCCGCCAGTGTGATGGATATCTGCAGAATTCCCC
CTTGGGGCGGGCGCTTCTGGGCGGACTGGAAAGACCGTCGTCTGTGGGTGACGGTGACGCCGATCGTGTTGATCACGTTTCCG
GCGGCGGTACAGGCGTACTTGTGGGAGCGGTATCGTCAACCGTGGGGTGCGACGGTGTGCGTGCTGGGTCTGTTGTTGGGTGA
GTGGGTCAACCGCTACTTCAACTTCTGGGGCTGGACGTATTTTCCGGTGGACTTCGTATTTCCGGCGATTTTGACGCCGGGTG
CGATTTTGCTGGACACGATGCTGTTGCTGAGCGGCAGCTATCTGTTCACGGCGATCATTGGCGGTTTGGGCTGGGGTTTGATT
TTCTACCCGGGCAACTGGCCGATCATTGCGCCGCTGCACGTACCGGTGGAGTACAGCGGGATGCTGATGTCGATTGCCGACAT
TCAGGGTTACAACTACGTGCGTACCGGAACCCCGGAATACATCCGGATGGTCGAGAAGGGCACGCTGCGTACGTTTGGTAAAG
ACGTTGCGCCGGAAGGGCGAATTCCAGCACACTGGCGGCCGTTACTAGTGGATCCGAGCTCGGTACCAAGCTTGGCGTAATCA
TGGTCATAGCCTGTTTCN

A DNA sequence trace can be seen in *Appendix 3*.

Sample 12c

TGGAAACAAAGNTTGACCATGATTACGCCAAGCTTGGTACCGAGCTCGGATCCACTAGTAACGGCCGCCAGTGTGCTGGAATT
CGCCCTTGGTGACTGGGACTTCTGGACTGACTGGAAAGACAGACGTCTGTGGGTAACCGTTTTACCGATTGTTGGTGTTACAT
TCCCTGCTGCCGTACAAGCCGTACTTTGGTATCGCTACCGTTTGCCGTTCGGTGCAGTTATTTCCGTTCTGGGTCTGCTGTTT
GGTGAGTGGGTTAACAGATACTTCAATTTCTGGGGATGGACTTATTTCCCAGTGAATTTCGTATTCCCATCACAATTCGTGCC
AGGCGCAATCGTTCTTGGTGTTATCCTGATGCTGTCCAACAGCATGCAATTAACAGCCGTTATCGGTGGTTTGGGCTATGGTC
TGTTGTTCTATCCTGGCAACTGGCCTGTCATCGCTCCATTGCACGTGCCTGTTGAATATAACGGTATGGTTATGACGCTGGCT
GACTTGCAAGGTTACCACTATGTAAGAACCGGTACCCCAGAGTACATCAGAATGGTTGAAAAAGGTACGTTGAGAACATTTGG
TAAAGACGTTGCGCCGGAAGGGCGAATTCTGCAGATATCCATCACACTGGCGGCCGCTCGAGCATGCATCTAGAGGGCCCANN
ANGNCAGTATAGNACANNNN

A DNA sequence trace can be seen in *Appendix 3*.

Sample 13a

GGCGACTGGGACTTCTGGTCCGACTGGAAAGACCGTCGTCTGTGGGTCACGGTGACCCCGATCGTGTTGGTCACCTTCCCGGC
TGCCGTTCAATACTTCCTGTGGGAACGGTTCCGTCAGCCCTGGGGTGCGACCGTGTGCGTACTGGCGCTGCTCTTCGGTGAAT
GGGTCAACCGCTACTTTAACTTCTGGGGCTGGACCTACTTCCCCGTGAACTTCGTGTTCCCGGCGGTCCTGGTTCCGGGCGCG
ATCCTGCTCGACGTTATCCTGATGCTCTCGGGCAGCTACCTGTTTGCTGCCATCATCGGCGGTCTGGCTTGGGGCCTGATTTT
CTATCCGGGCAACTGGCCGGTGATTGCTCCGCTGCACGTGCCGGTGGAATACAACGGCATGTTGATGTCCATTGCCGACATCC
AGGGCTACAACTATGTTCGTACCGGTACCCCCGAGTACATCCGCATGGTTGAGAAAGGCACCCTGCGTACCTTCGGTAAAGAC
GTTGCTCCGGAAGGGCGAATTCCAGCACACTGGCGGCCGTTACTAGTGGATCCGAGCTCGGTACCAAGCTTGGCGTAATCATG
GTCAANNCTGTTTTCNA

A DNA sequence trace can be seen in *Appendix 4*.

Sample 13b

GGGGACTGGGACTTCTGGACTGACTGGAAAGACCGTCGTCTGTGGGTAACCGTATTGCCAATCGTTGGTATTACATTCCCAGC
CGCTGTTCAAGCAGTTCTTTGGTACCGCTATCGTCTGCCATTCGGCGCAATGCTGGCTGTTCTTGGCCTGTTGTTCGGTGAAT
GGGTTAACAGATATTTCAACTTCTGGGGATGGACTTACTTCCCAGTTAACTTTGTATTCCCATCACAATTCGTTCCAGGCGCA
CTGGTTCTGGACGTAATTCTGATGCTGTCTAACAGCATGCAGTTGACTGCTGTTATCGGCGGTCTGGCATACGGCTTGTTGTT
CTATCCTGGCAACTGGCCTGTCATCGCTCCATTGCACGTGCCTGTTGAATACAACGGCATGGTAATGACGCTGGCTGACTTGC
AAGGTTACCACTATGTAAGAACTGGTACTCCAGAATACATCAGAATGGTTGAAAAAGGTACTTTGAGAACTTTCGGTAAGGAC
GTTGCTCCGGAAGGGCGAATTCCAGCACACTGGCGGCCGTTACTAGTGGATCCGAGCTCGGTACCAAGCTTGGCGTAATCATG
NNATNCTTATTTTCAA

A DNA sequence trace can be seen in *Appendix 4*.

Sample 13d

GGGGACTGGGACTTCTGGGCCTTCTTCTCCGCGTTCCGTTCGGCGCCGTTTTCGCGGCTCTGGGCCTTCTGATTGGCGAGTGG
ATTAACCGCTACGTCAACTTCTGGGGCTGGACGTATTTCCCGATCAGCCTCGTTTTCCCGTCTGCCCTGATCGTTCCGGCGAT
CTGGCTTGACGTGATCCTGCTCCTGTCGGGTTCCTATGTGATCACGGCGGTTGTCGGTTCGCTGGGCTGGGGTCTGCTGTTCT
ACCCGAACAACTGGCCGGCGATCGCCGCCTTCCACCAGGCGACCGAGCAGCATGGTCAGCTGATGACCCTCGCCGACCTGATT
GGTCTGCACTACGTCCGCACGTCGATGCCGGAATACATCCGCATGGTCGAGCGCGGCACGCTGCGCACGTTCGGTAAAGACGT
TGCTCCGGAAGGGCGAATTCCAGCACACTGGCGGCCGTTACTAGTGGATCCGAGCTCGGTACCAAGCTTGGCGTAATCATGGT
CATAGCTGATTCCAA

A DNA sequence trace can be seen in *Appendix 5*.

Sample 15b

GGAGACTGGGACTTCTGGGTTGACTGGAAAGACCGCCGTCTATGGGTGACCGTGGTGCCGATCGTGTTGGTCACCTTCCCGGC
CGCCACCCAGGTGATGCTGTGGGAGCGTCTGCGTCTTCCTTGGGGCGCGACGGTATGCGTATTGGCCCTGTTGTTTGGTGAGT
GGATCAACCGCTACTTCAACTTCTGGGGTTGGACCTACTTCCCGATCACCCTGTGCTTCCCCTCGCAGATCGTGCCGGGCGCC
ATCCTGCTCGACGTCGTTCTGTTGCTCTCCGGCAGCTACCTGCTCACCGCCATCGTGGGCGGCATGGCTTGGGGCTTGATCTT
CTACCCGGGCAACTGGCCGGTGATCGCGCCGTACCACGTGCCGGTGGAATACAACGGCATGCTGCTGTCGGTGGCCGACTTGC
TGGGCTACAACTATGTTCGTACCGGCACCCCCGAGTACATCCGCATGGTTGGGAAAGGCACCCTGCGTACCTTCGGTAAAGAC
GTTGCGCCGGAAGGGCGAATTCCAGCACACTGGCGGCCGTTACTAGTGGATCCGAGCTCGGTACCAAGCTTGGCGTAATCATG
GTCAANNTTTTTTTCAA

A DNA sequence trace can be seen in *Appendix 5*.

Sample 15d

GGTGACTGGGACTTCTGGACTGACTGGAAAGACCGTCGTCTGTGGGTAACCGTATTGCCAATCGTTGGTATTACATTCCCAGC
CGCTGTTCAAGCAGTTCTTTGGTACGGCTATCGTCTGCCATTCGGCGCAATGCTGGCTGTTCTTGGCCTGTTGTTCGGTGAAT
GGGTTAACAGATATTTCAACTTCTGGGGATGGACTTACTTCCCAGTTAACTTTGTATTCCCATCACAATTCGTTCCAGGCGCA
CTGGTTCTGGACGTAATTCTGATGCTGTCTAACAGCATGCAGTTGACTGCTGTTATCGGCGGTCTGGCATACGGCTTGTTGTT
CTATCCTGGCAACTGGCCTGTCATCGCTCCATTGCACGTGCCTGTTGAATACAACGGCATGGTAATGACGCTGGCTGACTTGC
AAGGTTACCACTATGTAAGAACTGGTACTCCAGAATACATCAGAATGGTTGAAAAAGGTACTTTGAGAACTTTCGGTAAGGAC
GTTGCGCCGGAAGGGCGAATTCCAGCACACTGGCGGCCGTTACTAGTGGATCCGAGCTCGGTACCAAGCTTGGCGTAATCATG
TNAANNTTTTTTTNCCNN

A DNA sequence trace can be seen in *Appendix 6*.

Sample 15e

GGAGACTGGGACTTCTGGGTTGACTGGAAGGATCGTCGTATGTGGCCGACGGTCGTGCCGATCCTTGGCGTGACCTTCTGCGC
GGCGTCGCAGGCTTTCTGGTGGGTTAACTTCCGCCTTCCGTTCGGCGCCGTGTTCGCGGCTCTGGGCCTCCTGATTGGCGAGT

GGATCAACCGCTACGTCAACTTCTGGGGCTGGACGTACTTCCCGATCAGCCTCGTGTTCCCGTCCGCTCTGATCGTTCCGGCG
ATCTGGCTCGACGTGATCCTGCTTCTGTCGGGTTCCTATGTGATCACGGCGGTTGTCGGTTCGCTGGGCTGGGGTCTGCTGTT
CTACCCGAACAACTGGCCGGCGATTGCGGCGTTCCACCAGGCGACCGAGCAGCATGGTCAGCTGATGACCCTGGCTGACCTGA
TCGGCCTTCACTGTGTCCGCACGTCGAAGCCGGAATACATCCGCATGGTCGAGCGCGGCACGCTGCGCACGTTCGGTAAAGAC
GTTGCTCCGGAAGGGCGAATTCTGCAGATATCCATCACACTGGCGGCCGCTCGAGCATGCATCTAGAGGGCCCCNATNAAACA
GNATAGNNGANNN

A DNA sequence trace can be seen in *Appendix 6.*

Sample 15i

GGTGACTGGGACTTCTGGGTTGACTGGAAGGATCGTCGTATGTGGCCGACGGTCGTGCCGATCCTGGGCGTGACCTTCTGCGC
GGCGTCGCAGGCTTTCTGGTGGGTTAACTTCCGTCTTCCGTTCGGCGCCGTTTTCGCGGCTCTGGGCCTGCTGATTGGCGAGT
GGATCAACCGCTACGTCAACTTCTGGGGCTGGACGTATTTCCCGATCAGCCTCGTGTCCCCGTCCGCTCTGATCGTTCCGGCG
ATCTGGCTCGACGTGATCCTGCTGCTGTCGGGCTCCTATGTGATCACGGCGGTTGTCGGTTCGCTGGGCTGGGGTCTGCTGTT
CTATCCGAACAGCTGGCCGGCGATTGCGGCGTTCCACCAGGCGACTGAGCAGCACGGCCAGCTGATGACGCTTGCTGACCTGA
TCGGCCTCCACTACGTCCGCACGTCGATGCCGGAATACATCCGCATGGTCGAGCGCGGCACGCTGCGCACGTTCGGTAAAGAC
GTTGCTCCGGAAGGGCGAATTCTGCAGATATCCATCACACTGGCGGCCGCTCGAGCATGCATCTAGAGGGCCNGAGNGNAACA
GCTAGANCGANNNN

A DNA sequence trace can be seen in *Appendix 7.*

Sample 16b

GGGGACTGGGACTTCTGGGTCGACTGGAAGGATCGTCGTATGTGGCCGACGGTCGTGCCGATTCTGGGCGTGACCTTCTGCGC
GGCTTCGCAGGCTTTCTGGTGGGTGAACTTCCGTCTGCCGTTCGGCGCGGTGTTTGCGGCTCTCGGCCTTCTGATTGGCGAGT
GGATCAACCGCTACGTCAACTTCTGGGGCTGGACCTACTTCCCGATCAGCCTCGTGTTCCCGTCTGCTCTGATCGTTCCGGCG
ATCTGGCTCGACGTGATCCTGCTTCTGTCGGGCTCCTATGTGATCACGGCGGTTGTCGGCTCGCTGGGCTGGGGTCTGCTGTT
CTATCCGAACAACTGGCCGGCGATCGCCGCCTTCCACCAGGCGACCGAGCAGCATGGTCAGCTGATGACGCTGGCCGACCTCA
TCGGCTTCCACTTCGTCCGCACCTCGATGCCGGAATACATCCGCATGGTCGAGCGCGGCACGCTGCGCACCTTCGGTAAAGAC
GTTGCGCCGGAAGGGCGAATTCTGCAGATATCCATCACACTGGCGGCCGCTCGAGCATGCATCTAGAGGGCCCANNNANNNNN
CTCGTNCTANNNNN

A DNA sequence trace can be seen in *Appendix 7.*

Sample 16c

GGGGACTGGGACTTCTGGGTTGACTGGAAGGATCGTCGTATGTGGCCGACGGTCGTGCCGATTCTTGGCGTGACCTTCTGCGC
GGCGTCGCAGGCTTTCTGGTGGGTGAACTTCCGTCTGCCGTTCGGCGCGGTGTTCGCGGCTCTCGGCCTGCTGATCGGCGAGT
GGATCAACCGCTACGTCAACTTCTGGGGCTGGACCTACTTCCCGATCAGCCTGGTGTTCCCGTCTGCTCTGATCGTTCCGGCG
ATCTGGCTGGACGTGATCCTGCTGCTGTCGGGCTCCTATGTGATCACGGCGGTTGTCGGTTCGCTGGGCTGGGGTCTGCTGTT
CTATCCGAACAACTGGCCGGCGATCGCCGCCTTCCACCAGGCGACCGAGCAGCATGGTCAGCTGATGACGCTTGCTGACCTGA
TCGGCCTCCACTACGTCCGCACGTCGATGCCGGAATACATCCGCATGGTCGAGCGCGGCACGCTGCGCACGTTCGGTAAAGAC
GTTGCGCCGGAAGGGCGAATTCTGCAGATATCCATCACACTGGCGGCCGCTCGAGCATGCATCTAGAGGGCCCNNNNNACAGC
TAGTAGANNNN

A DNA sequence trace can be seen in *Appendix 8.*

Sample 16e

GGGGACTGGGACTTCTGGACGGACTGGAAAGACAGACGTCTGTGGGTAACGGTTGTGCCGATCGTAGCTGTTACTTTCCCTGC
GGCTGTGCAAGCTTTCGCATGGGGTCGTTACCGTTTGCCTTGGGGCGCAACCGTTTGCGTTCTGGGCTTGCTGTTCGGTGAAT
GGGTTAACAGATACTTCAACTTCTGGGGTTGGACATACTTCCCGGTTAACTTCGTATTCCCATCACAATTGATTCCTAGCGCT
ATCCTGCTGGACGTGGTTCTGTTGCTGTCCAACAGCTACACCTTCACCGCTGTTGCTGGCGCTATGGGTTGGGGCTTGATTTT
CTATCCTAGCAACTGGCCGGTTATTGGTCCATTACACGTGCCTGTTGAATATAACGGCATGATGATGACTTTGGCTGACTTAC
AAGGTTACCACTATGTAAGAACCGGTACTCCTGAATACATCCGTATGGTTGAAAAAGGTACATTGAGAACTTTCGGTAAAGAC
GTTGCTCCGGAAGGGCGAATTCCAGCACACTGGCGGCCGTTACTAGTGGATCCGAGCTCGGTACCAAGCTTGGCGTAATCATG
GTCATAGCTTGTTTCCANN

A DNA sequence trace can be seen in *Appendix 8.*

The 16 sequences were then aligned together with type I and II methanotrophs in the sequence alignment program “MUSCLE”. The type I methanotrophs were Methylobacteralbus and Methylomonasmethanica. The type II methanotrophs were Methylocystis and Methylosinus. The sequences corresponds to the pmoA gene and the sequences can be found in *Appendix 9.*

The sequences were then put in a program called “HIV sequence database” which convert the sequences to file formats suitable for making a phylogenetic tree. A phylogenetic tree, based on maximum likelihood was then made with the program “PhyML”.

The phylogenetic tree made from the 16 samples, plus pmoA sequences from two type I methanotrophs and from two from type II, can be seen in *figure 9*, in form of a phylogram. A cladogram with bootstrap values can be seen in *figure 10.* A sequence alignment with the program MUSCLE can be seen in *Appendix 10.* The phylogenetic tree in Newick format can be seen in *Appendix 11*. A slanted cladogram with bootstrap values can also be seen in *Appendix 11.*

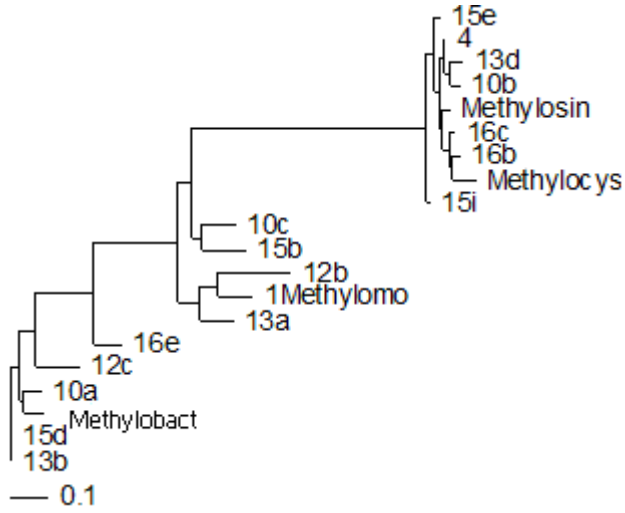


Figure 9. A phylogram based on the pmoA gene from the samples plus from known methanotrophs, both type I and II (The bottom scale measures the genetic distance in substitutions per nucleotide).

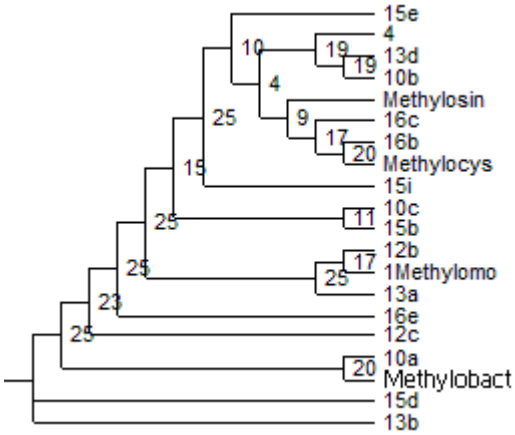


Figure 10. A rectanuglar cladogram with bootstrap values, based on the pmoA gene from the samples plus from known methanotrophs, both type I and II. 25 bootstrap data sets.

Another phylogenetic tree was made, based on all 16 samples together with pmoA sequences from *Methylobacter*, *Methylomicrobium*, *Methylomonas*, *Methylocaldum*, *Methylothermus*, *Methylosarcina*, *Methylohalobius*, *Methylosoma*, and *Methylococcu*s (these belong to type I methanotrophs) and *Methylocystis*, *Methylosinus* and *Methylocapsa* (these belong to type II methanotrophs).The sequences of the pmoA gene can be found in *Appendix 12*. A phylogram of above sequences can be seen as a phylogram in *figure 11* and as a cladogram with

bootstrap values in *figure 12*. A sequence alignment from MUSCLE can be seen in *Appendix 13.* The phylogenetic tree in Newick format can be seen in *Appendix 14.*

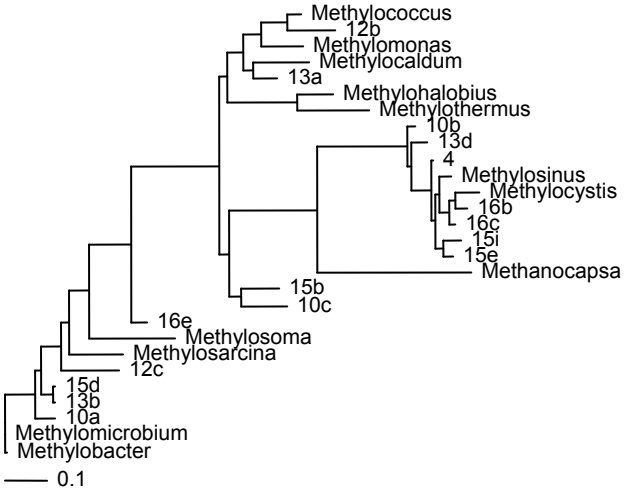


Figure 11. A phylogram based on the pmoA gene from the 16 samples together with 12 known methanotrophs, both type I and II (The bottom scale measures the genetic distance in substitutions per nucleotide).

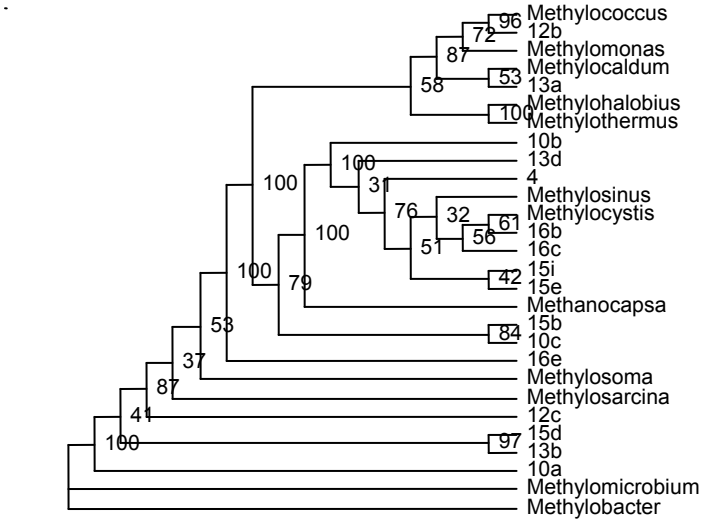


Figure 12. A rectanuglar cladogram with bootstrap values, based on the pmoA gene from the 16 samples together with 12 known methanotrophs, both type I and II. 100 bootstrap data sets.

Discussion

The DNA extraction of the soil samples worked good with the FastDNA® SPIN Kit for Soil, which can be seen in *figure 3* in Material and Methods. The other method, described in Griffith et al., 2000, did not work, only for two samples which is not much. Probably some concentrations had to be changed to suit this kind of soil samples. The PCR for the FastDNA® SPIN Kit for Soil samples worked well for most of the reactions. The cloning and transformation did not work at first, which can be seen in *figure 7*. There was no PCR product

amplified which means that either the DNA did not get ligated into the vector or the vector did not get inserted to the cells. It worked well when the PCR products were fresh, which can be seen in *figure 8*. The cloning of the control samples can also indicate that, because here he PCR products were fresh, which can be seen in *figure 5*. Antoher misstake could have been that the competent *E.* coli cells were old, or had laid at the top draw in the freezer and therefore were not as active as they should. The *E*. coli cells were then replaced with new fresh ones.

The 16 samples sent to sequencing were as mentioned earlier; 4, 10a, 10b, 10c, 12b, 12c, 13a, 13b, 13d, 15b, 15d, 15e, 15i, 16b, 16c and 16e. Sample 4 contained soil samples 19, 20, 21 and 22 which comes from a square treated with only rice, which can be seen in *table 1* in Introduction. Sample 10a, 10b and 10c, contained the soil samples 75, 76, 77 and 78 that comes from a square treated with rice and mungbean. Sample 12b and 12c contained the soil samples 91, 92, 93 and 94 that comes from a square treated with rice and mungbean. Sample 13b and 13d contained the soil samples 99, 100, 101 and 102 that comes from a square treated with rice, mungbean and maize. Sample 15b, 15d, 15e and 15i contained the soil samples 115, 116, 117 and 118 that comes from a square treated with rice, mungbean and maize. Sample 16b, 16c and 16e contained the soil samples 123, 124, 125 and 126 that comes from a square treated with rice, mungbean and maize.

So there is one sample with methanotrophs that comes from a soil square treated with only rice. Two samples comes from soil treated with rice and mungbean and tree sample that comes from soil treated with rice, mungbean and mize. So it is not clear yet wether the methanotrophs in the soil are affected by these three types of treatments. More tests have to be done to see any clear statistical difference because even though differences can be seen here, there can be errors like spreading of plants across the squares borders. The operational taxonomic units, OTUs, in the phylogenetic tree are all samples, the hypothetical taxonomic units, HTUs, can not be told.

There are only one square with the treatment ‘only rice’ that contained methanotrophs. A conclusion could be that there is too much methane in the soil for the methanotrophs to be pleased and able to grow. Maybe the soil is exposed to too intensive cultivation, the amount of methane produced will increase because the methanotrophs have no longer the ability to oxidate enough. Since methane absorbs heat radiation better than carbon dioxide, this can lead to a big threat to the global warming.

By looking at the phylogenetic tree, the clustres can easiest be seen in the phylogram, in *figure 9* in Results. There are two subgroups. The samples 4, 10b, 13d, 15e, 15i, 16b and 16c will form a cluster together with the type II methanotrophs, Methylocystis and Methylosinus. The samples 10c, 12b, 13a and 15b form a cluster with the type I methanotroph, Methylomonas. Rest of the samples, 10a, 12c, 13b, 15d and 16e will form a cluster with the type I methanotrophs Methylobacter. Sample 10a are very closely related to Methylobacter, 12b to Methylomonas, 16b to Methylocystis. Sample 16b plus 16c are also closely related to Methylosinus. This can also be indicated by the bootstrap values, that all are 20 of 25, which can be seen in *figure* 10. The genera that every sample belongs to, can not be determined for all samples. But what can be stated is that both typ I and type II methanotrophs exists in the soil. When there is a mixed culture of type I and II methanotrophs, tests have been shown that metabolic differences most often leads to competition and results in occupation of different niches by the two groups. But on the other hand, these two types should complement

eachother, while type I is favoured in environment with low concentrations of methane and the opposite for type II (Macalady JL et al., 2002).

Another thing worth mentioning is the sample 16e which has its own branch between the two clustres of type I methanotrophs, it can easiest be seen in the phylogram, *figure 7.* There is a third group of methanotrophs called type X. Type X belongs to the gamma subdivision of the Proteobacteria, just as type I (Hanson RS et al., 1996). So type X has phylogenetic similar to type I but some metabolic attributes similar to type II (Macalady JL et al., 2002). So further interesting analysis would be testing of the samples metabolic attributes. For example, type X and type II methanotrophs, have both the ability to fix nitrogen. Whereas type I methanotrophs have not, so the fixation of nitrogen can be tested to distinguish it as type X.

In the second phylogenetic tree, *figure 11* and 12, Methylosphaera was the only class of methanotrophs type I missing in the phylogenetic tree and *Methylocella* from type II. This is because the sequence for these pmoA gene could not be found in a sequence database. In this phylogenetic tree, a more define structure can be seen. Sample 12b is closely related to *Methylococcus*, 13a to *Methylocaldum* and 12c to *Methylosarcina*. The samples 10b, 13d, 4, 16b, 16c, 15i, 15e seem to belong to the type II methanotrophs. This can also be verified by looking at the bootstrap values in the cladogram, *figure 12.* Rest of the samples, 15b, 10c, 16e, 15d, 13b and 10a are difficult to put in a subgroup. Although, sample 15d and 13b are very closely related.

Sample 4 came from a square treated with only rice, which can be seen in *table 1* in Introduction. Sample 4 belongs to the type II methanotrophs. Sample 10a, 10b and 10c come from a square treated with rice and mungbean and can be found in clusters with both type I and II methanotrophs. The same is for sample 13a, 13b and 13d but they comes from a square treated with rice, mungbean and maize.Sample 12b and 12c come from a square treated with rice and mungbean and can be found in cluster with type I methanotrophs. Sample 15b, 15d, 15e and 15i come from a square treated with rice, mungbean and maize and forms clusters with both types. Sample 16b, 16c and 16e come from a square treated with rice, mungbean and maize and also forms clusters with both types of methanotrophs. So the samples are equally distributed over the two types of methanotrophs. By looking at the second phylogenetic tree, there are more samples than 16e that can belong to type X. Samples 15d, 13b and 10a form an own subgroup in the cluster with type I methanotrophs. Sample 15b and 10c form a cluster between the two types and seems to be related. Comparison between the two phylogenetic trees indicate how easy it is to make errors. For example, in the first phylogenetic tree, *figure 9*, sample 4 seems to belong to type II methanotrophs but in the secons phylogenetic tree, *figure 11*, it looks like it is realted with type I methanotrophs. So, more samples have to be included, bothknown and unknown values, for further phylogenetic and metabolic studies to make statistical reliable values.

Acknowledgement

I will like to thank my wonderful supervisor Nils Högberg for pedagogical supervision. A special thanks to Xuan Do Thi for her soil samples that I used in this project and also a thank to all the helpful coworkers in the lab like Maria Jonsson and Rena Gadjieva.

References

Articles

Brigmon, R. L. 2001. *Methanotrophic bacteria: use in bioremediation*, *In*, G. B. Bitton (ed.), Encyclopedia of environmental microbiology.

Conway de Macario E, Macario AJ. *Methanogenic archaea in health and disease: a novel paradigm of microbial pathogenesis.* Int J Med Microbiol. 2009 Feb;299(2):99-108.

Don R H, Cox P T, Wainwright B J, Baker K, and Mattick J S. *'Touchdown' PCR to circumvent spurious priming during gene amplification.* Nucleic Acids Res. 1991 July 25; 19(14): 4008.

Griffiths RI, Whiteley AS, O'Donnell AG, Bailey MJ. *Rapid method for coextraction of DNA and RNA from natural environments for analysis of ribosomal DNA- and rRNA-based microbial community composition*. Appl Environ Microbiol. 2000 Dec;66(12):5488-91.

Hanson RS and Hanson TE**,** *Methanotrophic bacteria*, Microbiol. Rev., Jun 1996, 439-471, Vol 60, No. 2

Hellström, D., Edström, M., Nordberg, Å., Olsson, L., (2001). *Projekt Hammarby Sjöstad, Anaerob behandling av avloppsvatten: Förstudie- Processrapport.* Stockholm Vatten.

Horz HP, Rich V, Avrahami S, Bohannan BJ. *Methane-oxidizing bacteria in a California upland grassland soil: diversity and response to simulated global change.* Appl Environ Microbiol. 2005 May;71(5):2642-52.

Knief C, Vanitchung S, Harvey NW, Conrad R, Dunfield PF, Chidthaisong A. *Diversity of methanotrophic bacteria in tropical upland soils under different land uses.* Appl Environ Microbiol. 2005 Jul;71(7):3826-31

Macalady J L, McMillan AMS, Dickens AF, Tyler SC and Scow KM, Population *dynamics of type I and II methanotrophic bacteria in rice soils*, Environmental Microbiology (2002) 4(3), 148–157

McDonald IR, Bodrossy L, Chen Y, Murrell JC. *Molecular ecology techniques for the study of aerobic methanotrophs.* Appl Environ Microbiol. 2008 Mar;74(5):1305-15. Epub 2007 Dec 28.

Rodhe,H. och Svensson, B.. *Impact on the greenhouse effect of peat mining and combustion.* Ambio 1995, 24: 221–225.

Salemi, M, Vandamme AM, *The phylogenetic handbook – A practical approach to DNA and protein phylogeny*, Cambridge University press 2003

Tchawa Yimga M, Dunfield PF, Ricke P, Heyer J, Liesack W. *Wide distribution of a novel pmoA-like gene copy among type II methanotrophs, and its expression in Methylocystis strain SC2*. Appl Environ Microbiol. 2003 Sep; 69(9):5593-602.

Internet pages

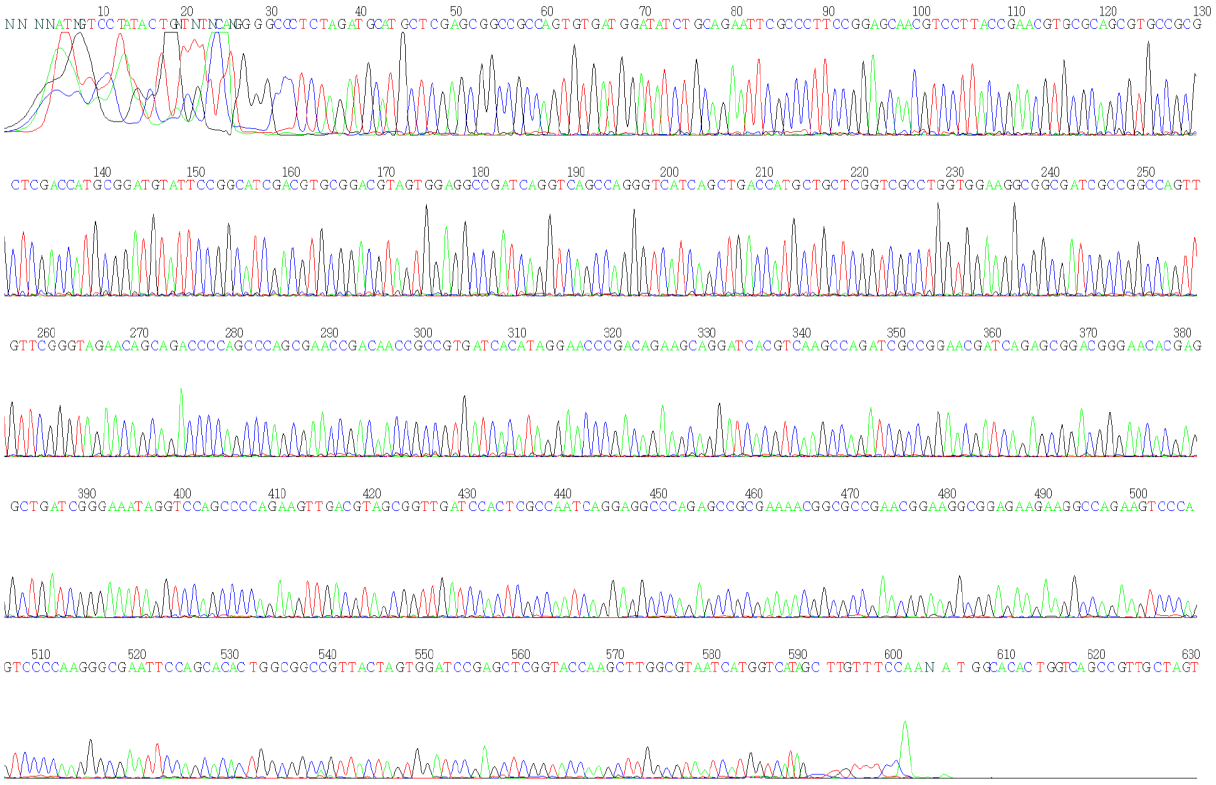
www.pcrstation.com/nested.pcr (2009-04-22)
www.invitrogen.com (2009-05-09)
www.ebi.ac.uk/Tools/clustalw2/index.html (2009-05-11)
www.hiv.lanl.gov/content/sequence/FORMAT_CONVERSION/form.html (2009-05-11)
www.rcsb.org/pdb/home/home.do, PDB id; 1yew (2009-14-5)

Lectures

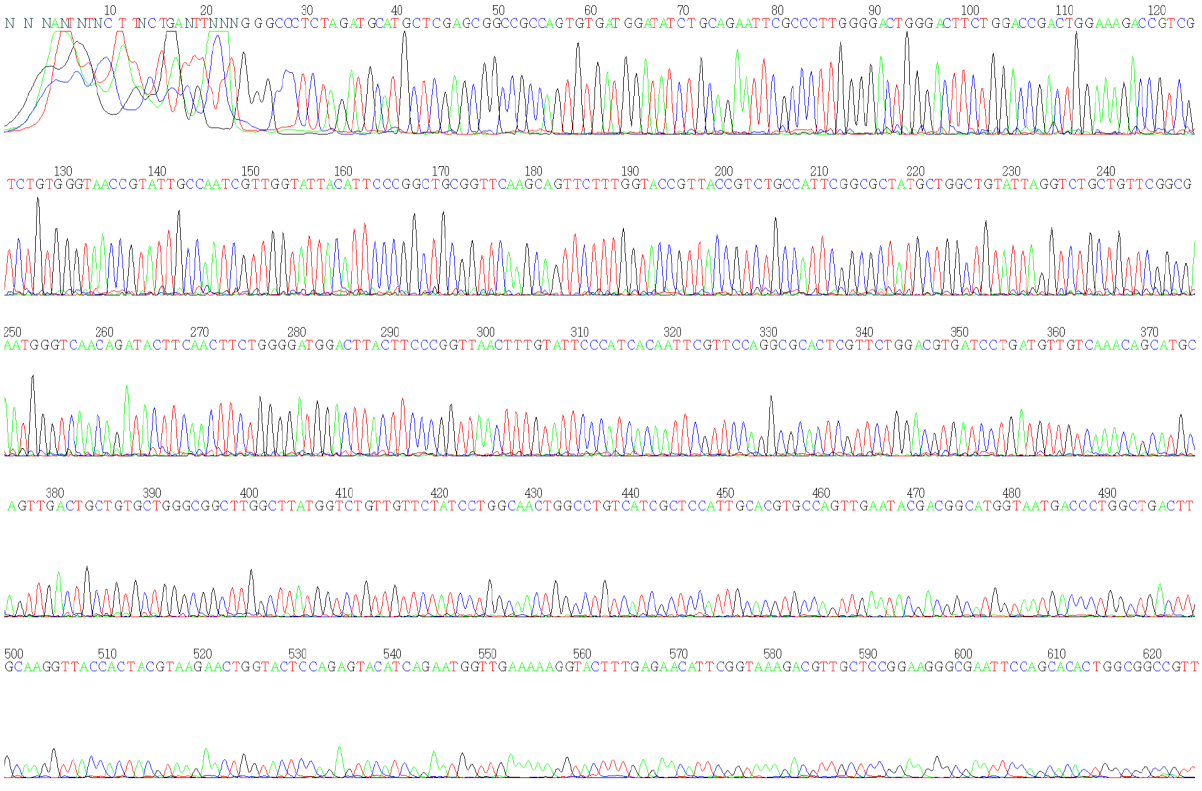
Anna Schnürer, *Biogasproduktion -från urtid till nutid, Institutionen för Mikrobiologi*, SLU ”Avloppsdag”, 27 November 2008

Appendix 1

Sample 4

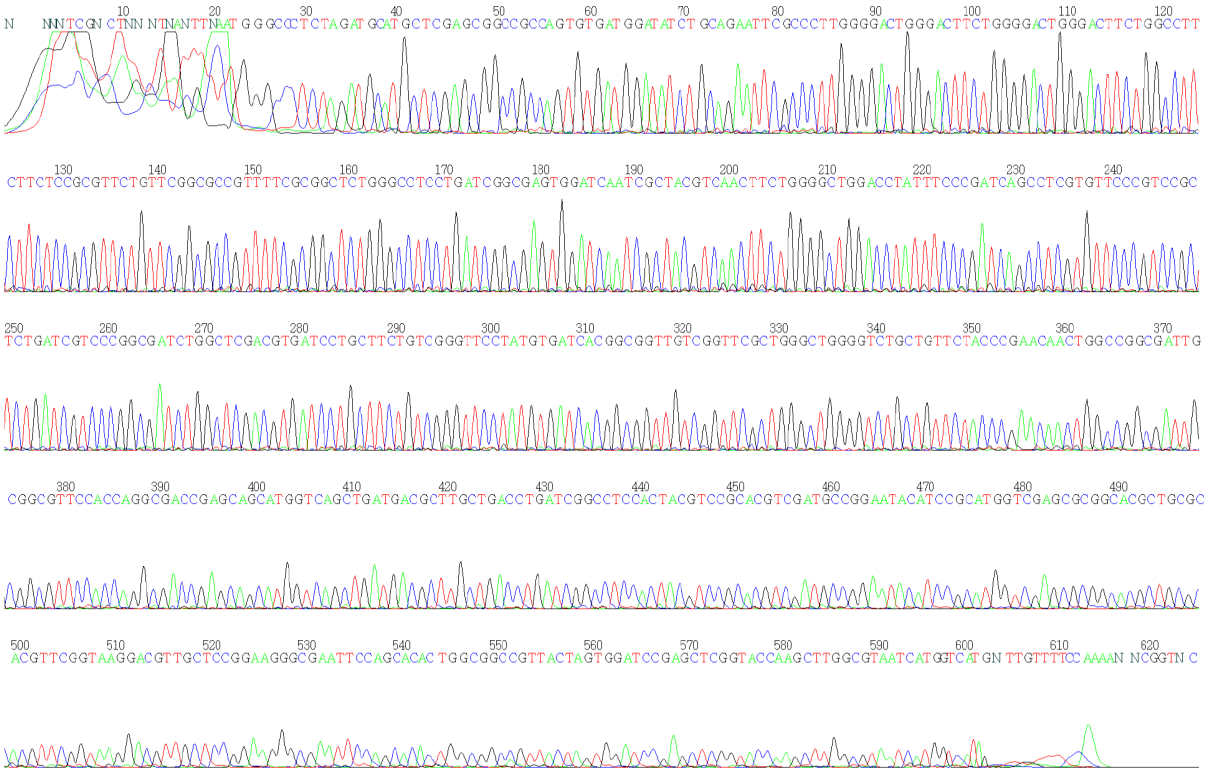


Sample 10a

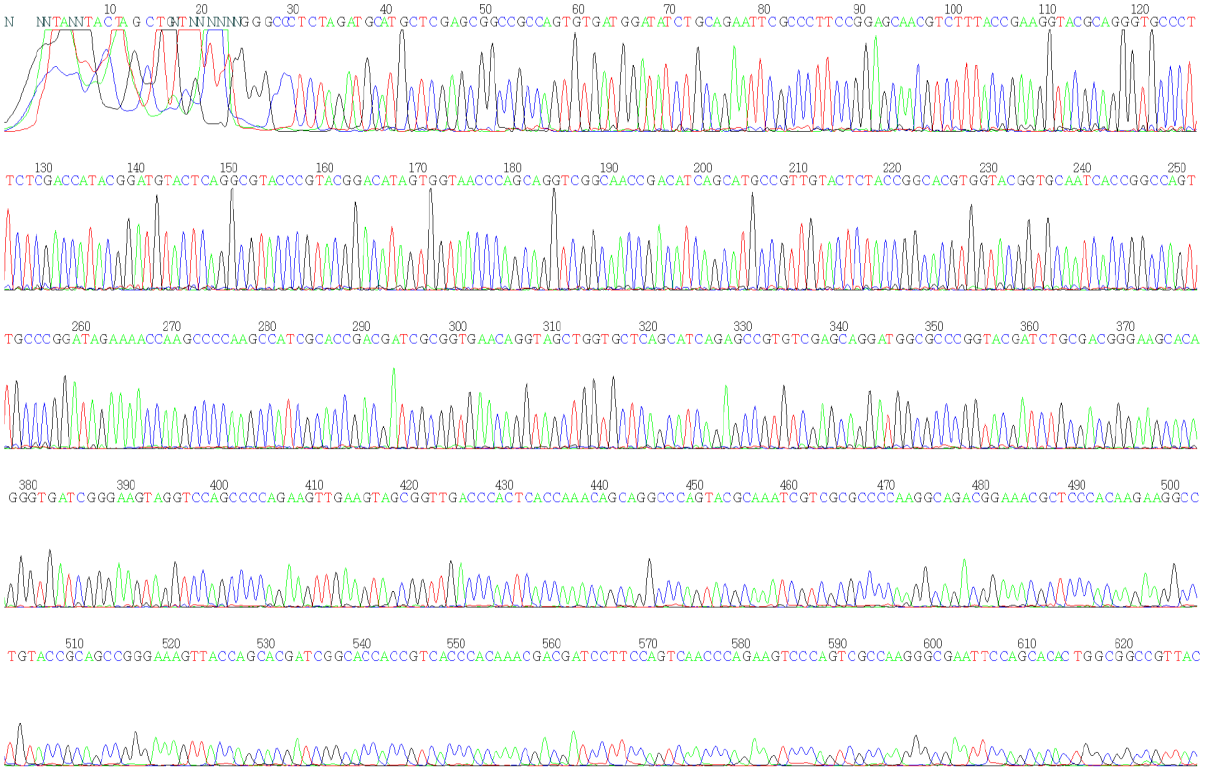


Appendix 2

Sample 10b

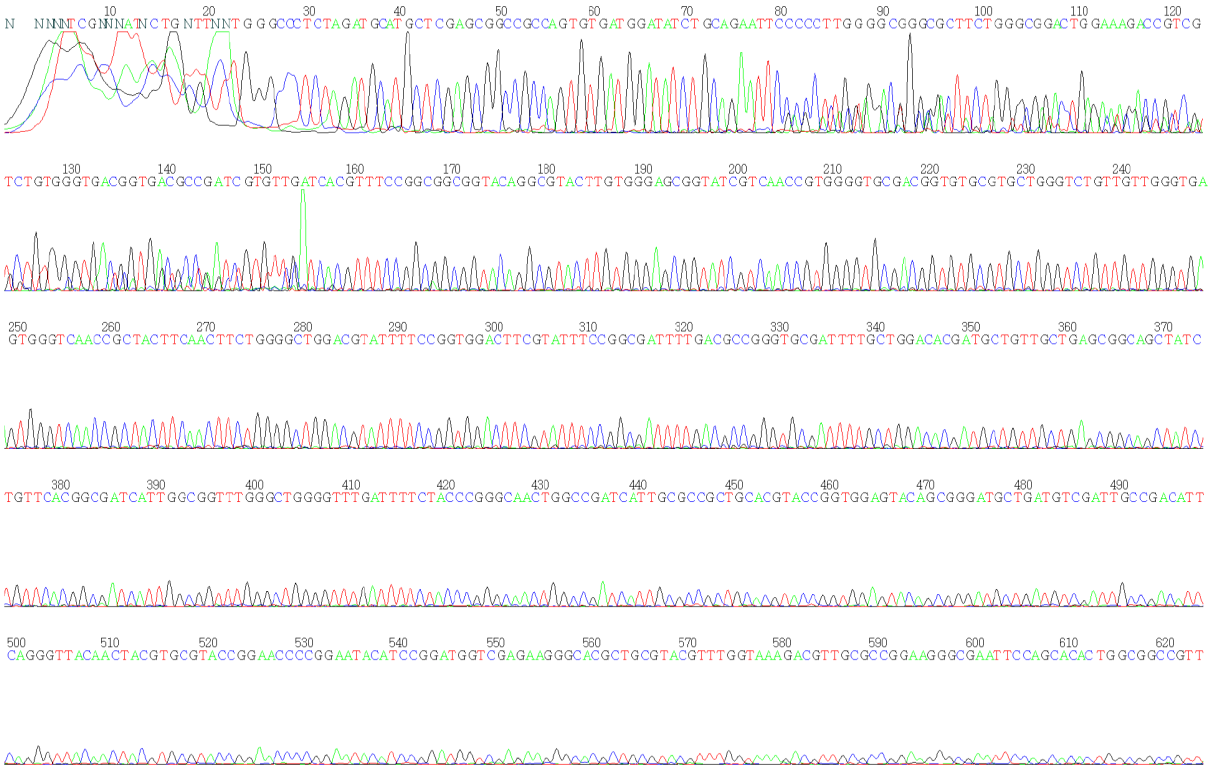


Sample 10c

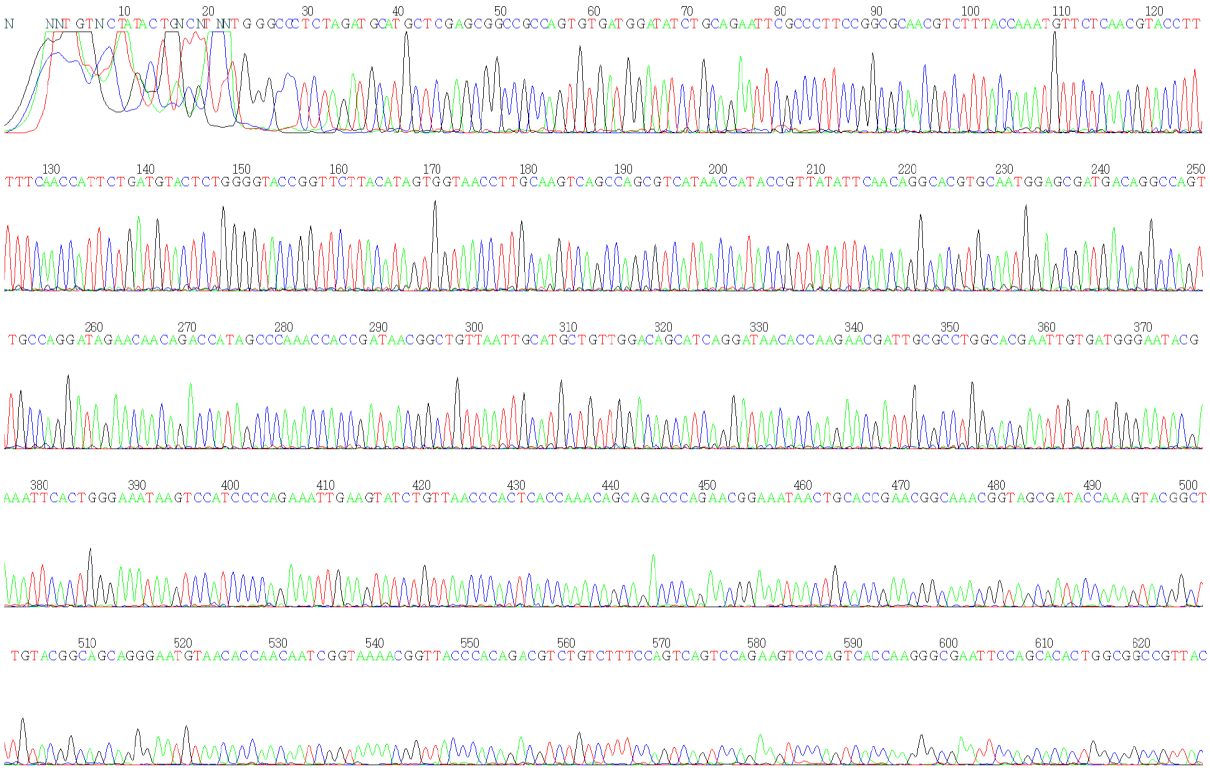


Appendix 3

Sample 12b

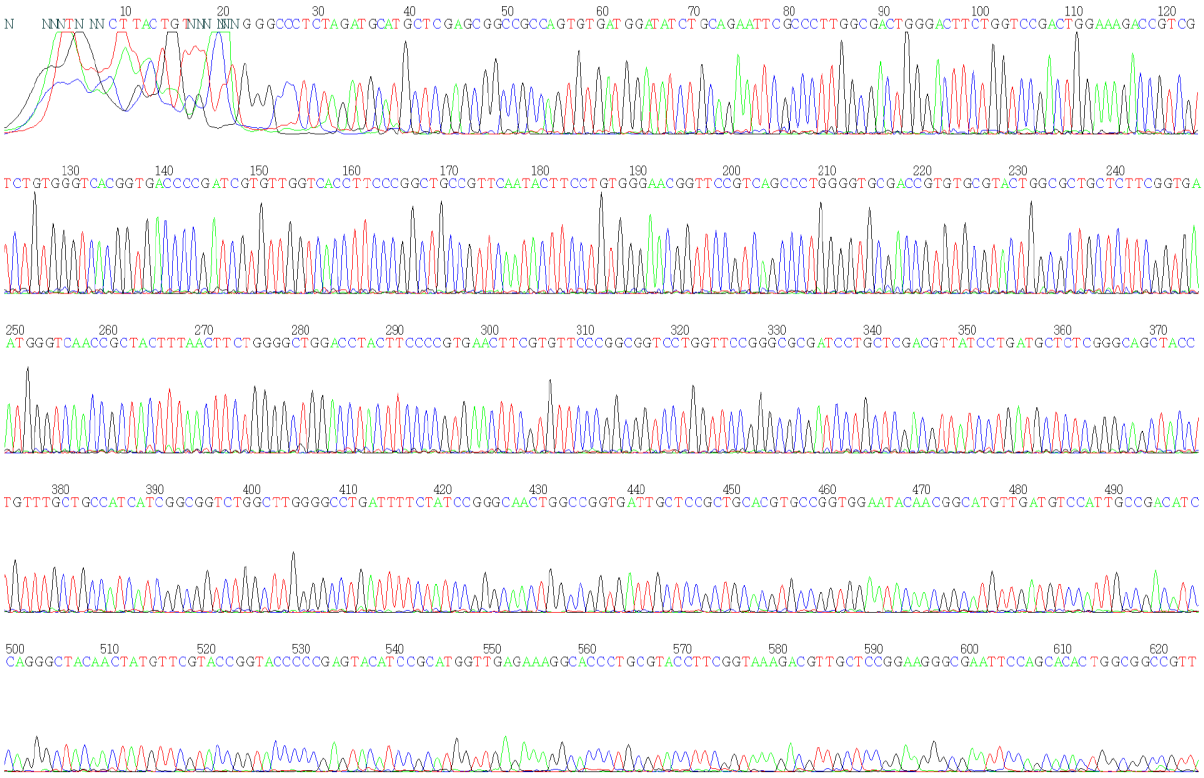


Sample 12c

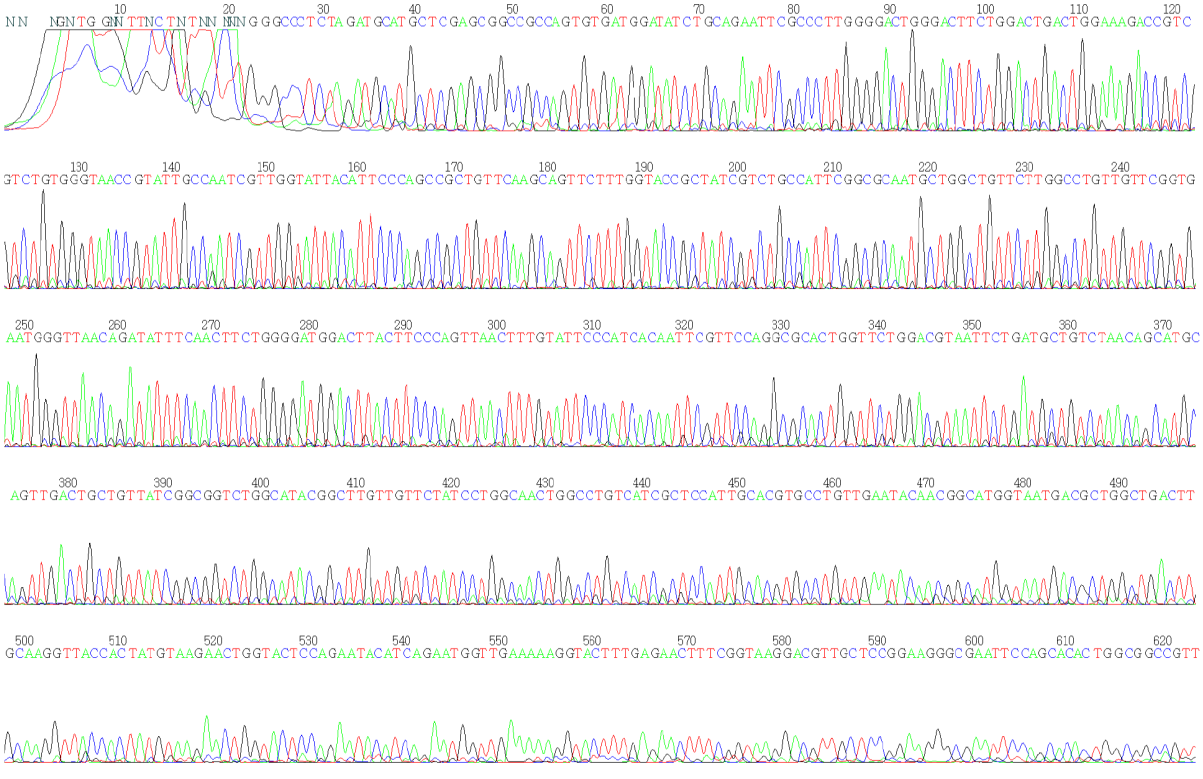


Appendix 4

Sample 13a

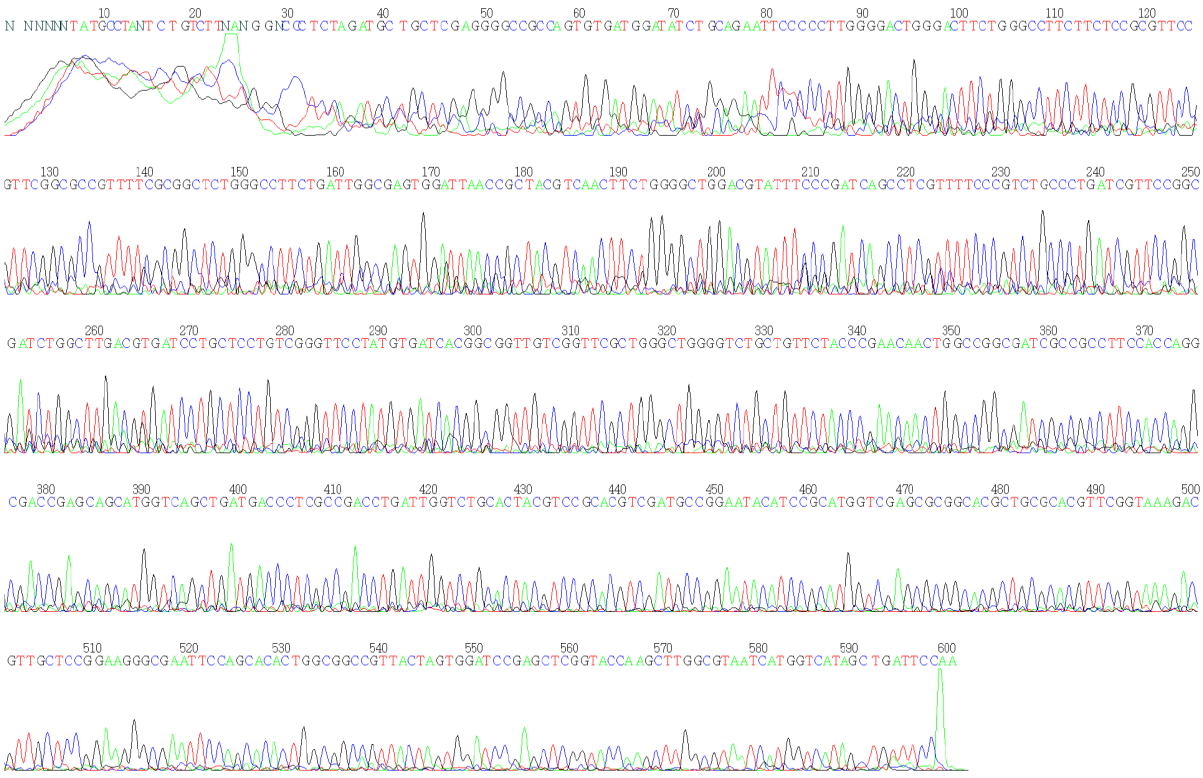


Sample 13b

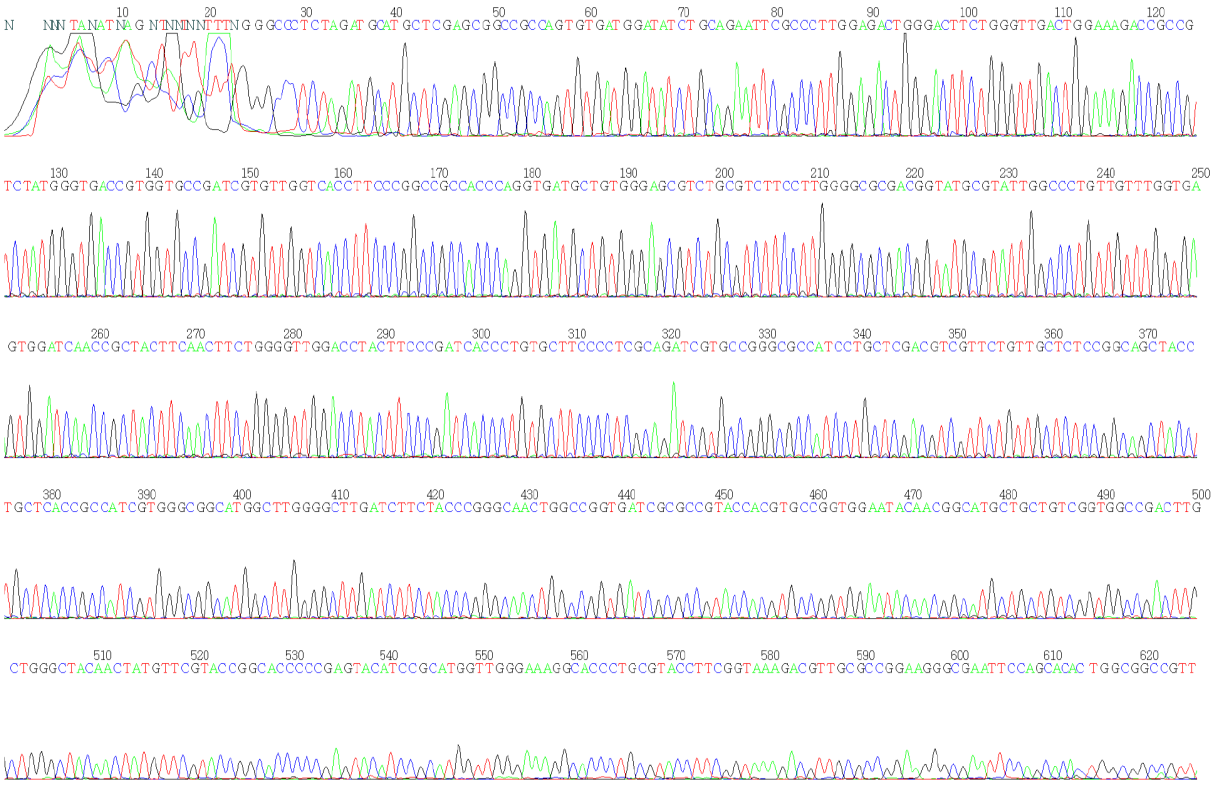


Appendix 5

Sample 13d

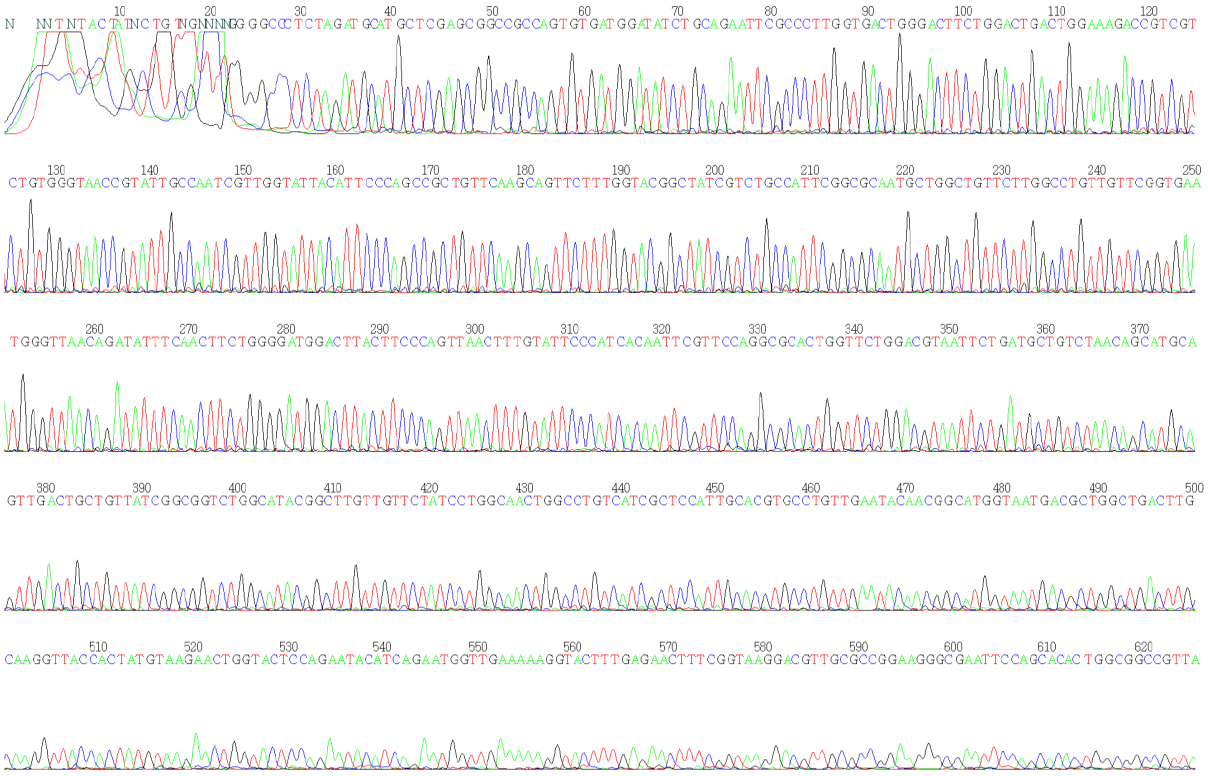


Sample 15b

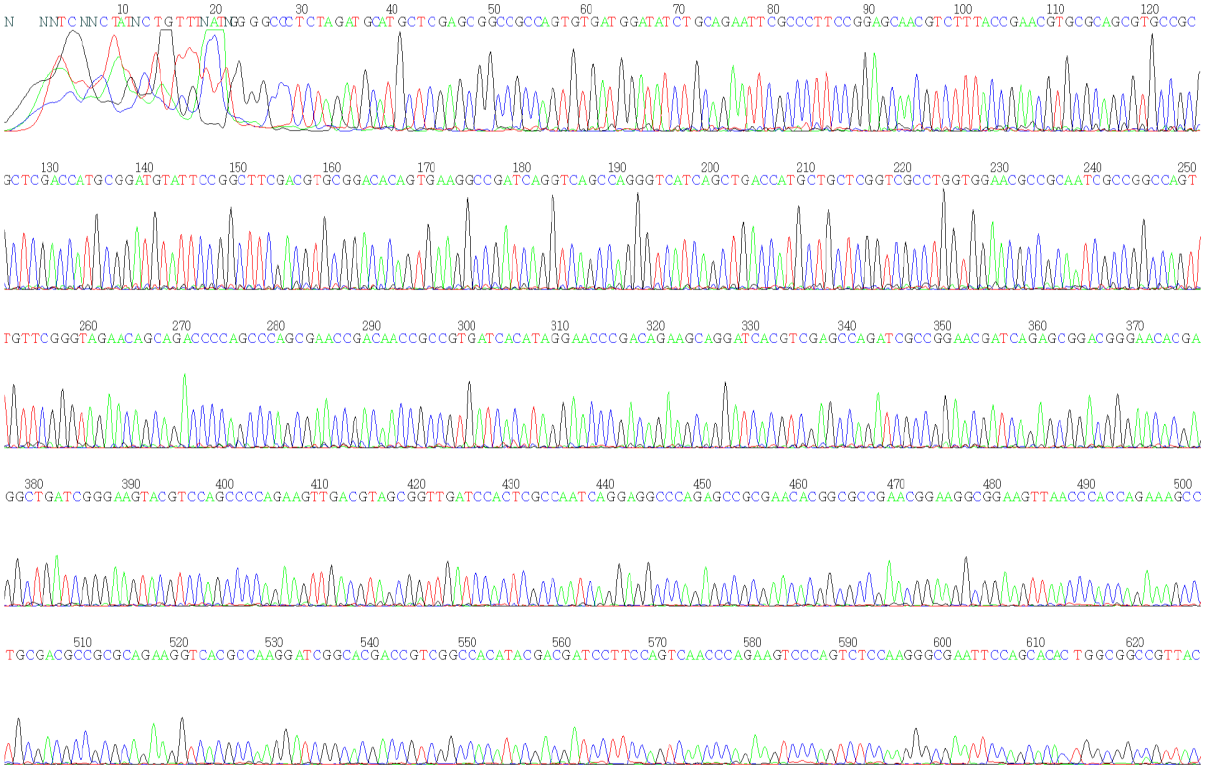


Appendix 6

Sample 15d

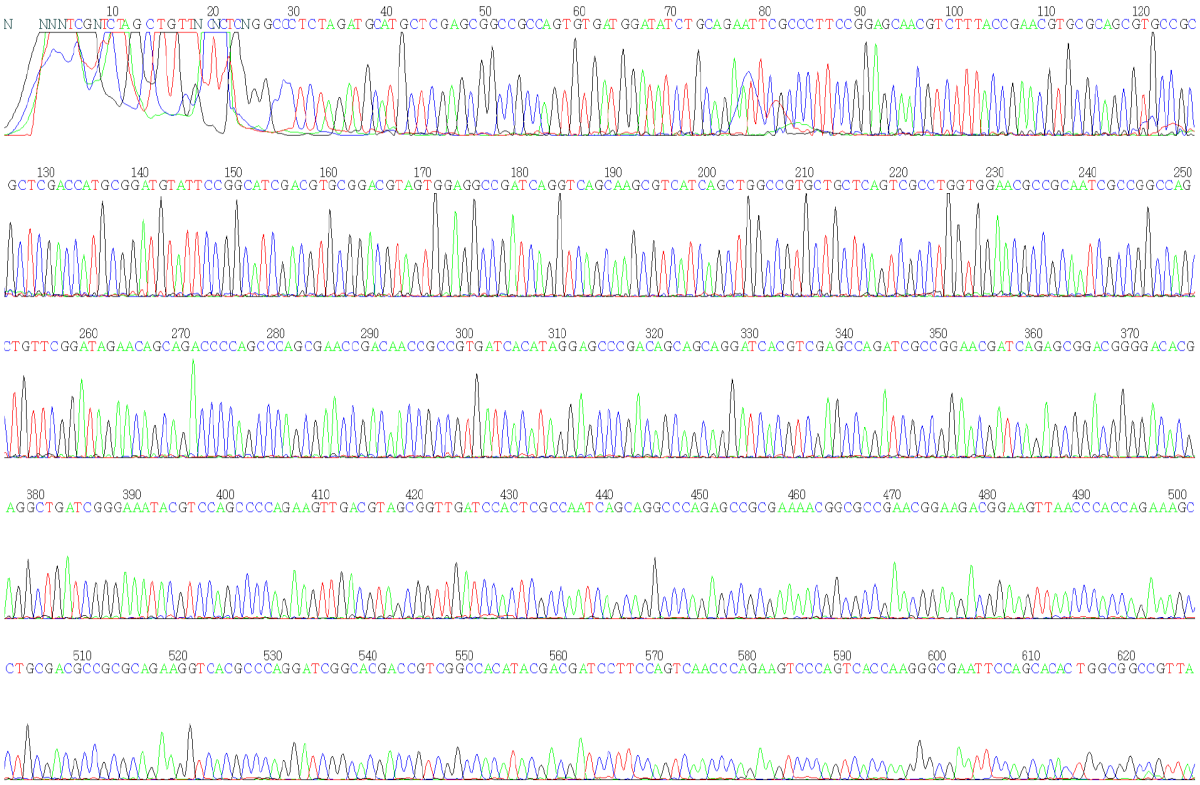


Sample 15e

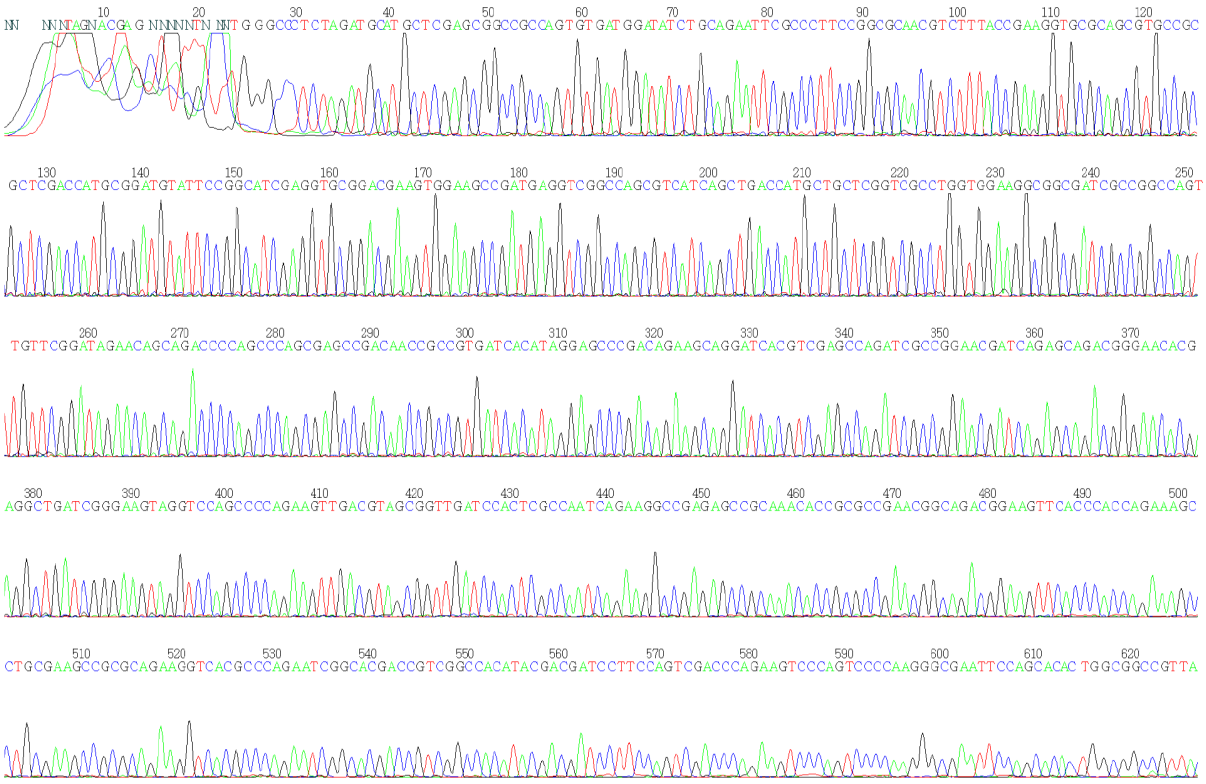


Appendix 7

Sample 15i



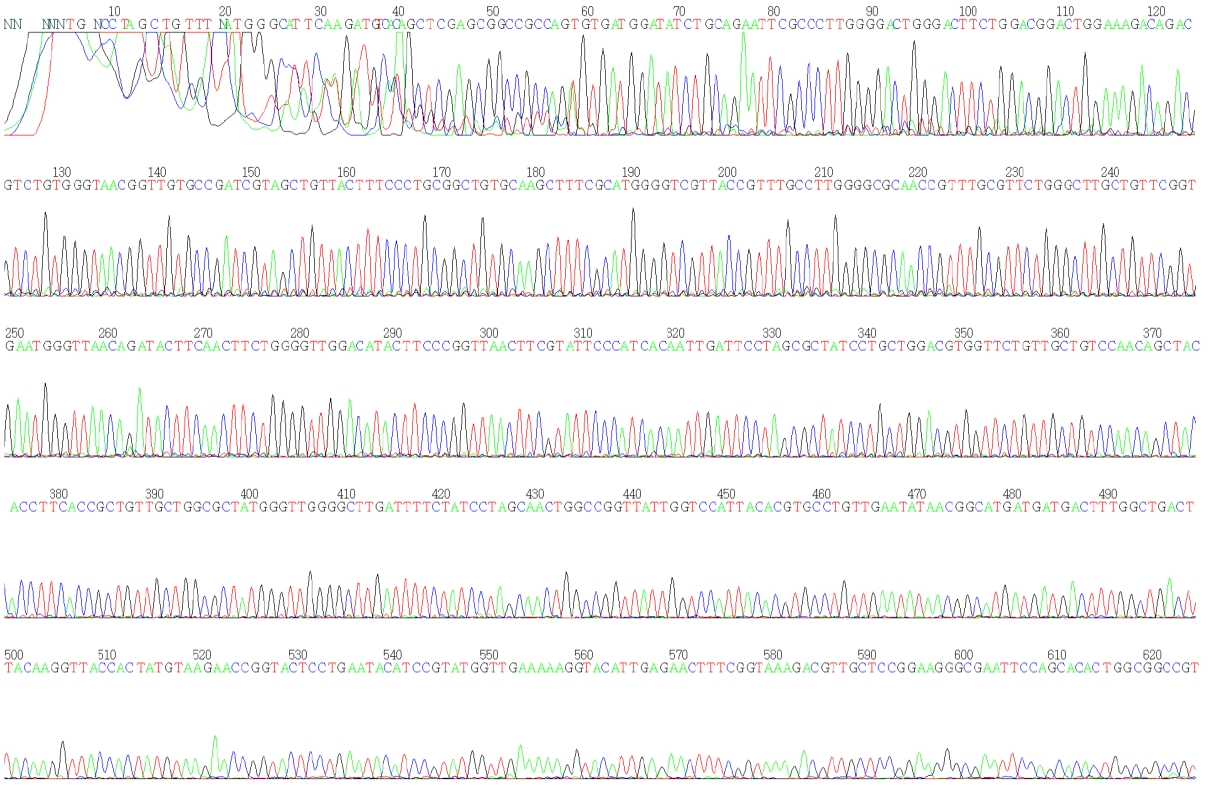
Sample 16b



Appendix 8

Sample 16c

Sample 16e



Appendix 9

Type I methanotrophs (pmoA gene)

Methylobacter

ACCGATTGGAAAGACCGTCGTCTGTGGGTAACCGTATTGCCAATCGTTGGTATTACTTTCCCAGCCGCTGTTCAAGCAGTTGT
GTGGTATCGTTGGCGTCTGCCATTCGGCGCAATGCTGGCTGTTCTGGGCCTGCTGTTCGGCGAATGGGTTAACAGATATTTCA
ACTTCTGGGGATGGACTTACTTCCCAGTTAACTTCGTGTTCCCATCACAATTCGCTCCAGGCGCAATCGTTCTGGACGTAATT
CTGATGCTGTCTAACAGCATGCAGTTGACTGCGGTTCTGGGCGGCTTGGCTTATGGGCTGTTGTTCTATCCTGGCAACTGGCC
GGTCATCGCTCCATTGCACGTGCCAGTTGAATACAACGGGATGGTAATGACCCTGGCTGACTTGCAAGGTTACCACTATGTAA
GAACCGGTACTCCAGAATATATCCGGATGGTTGAAAAAGGTACTCTGAGAACTTTCGGTAAGGACGTTGCTCCAGTATCC

Methylomonas

GACTGGGACTTCTGGTCGGACTGGAAAGACCGTCGACTGTGGGTCACGGTGACCCCGATCGTACTGGTCACCTTCCCGGCGGC
CGTACAATCCTACCTGTGGGAGCGGTATCGTCTGCCCTGGGGAGCCACCGTGTGCGTCCTGGGTCTGCTGCTGGGCGAGTGGA
TCAACCGTTATTTCAACTTCTGGGGCTGGACCTACTTCCCGATCAACTTCGTGTTCCCTGCCTCGCTGGTGCCGGGCGCCATC
ATCCTGGACACCGTGCTGATGCTGTCGGGCAGCTACCTGTTCACCGCGATCGTCGGTGCGATGGGCTGGGGTCTGACCTTCTA
CCCGGGCAACTGGCCGATCATCGCGCCGCTGCACGTGCCGGTGGAATACAACGGCATGCTGATGTCGATCGCCGACATCCAGG
GTTACAACTATGTGCGTACGGGTACGCCTGAGTACATCCGCATGGTAGAGAAGGGCACCCTGCGTACCTTCGGTAAGGAC

Type II methanotrophs (pMOA gene)

Methylocystis

GTTGACTGGAAGGATCGTCGTATGTGGCCGACGGTCGTGCCGATTCTCGGCGTGACCTTCTGCGCGGCGTCGCAGGCGTTCTG
GTGGGTGAACTTCCGTCTGCCGTTTGGCGCGGTGTTCGCGGCTCTTGGCCTTCTGATTGGCGAGTGGATCAACCGCTACGTCA
ACTTCTGGGGCTGGACCTACTTCCCGATCAGCCTTGTGTTCCCGTCTGCGTTGATCGTTCCGGCGATCTGGCTTGACGTGATC
CTGCTTCTGTCGGGGTCCTATGTGATCACGGCGATTGTTGGTTCGCTGGGCTGGGGTCTGTTGTTCTACCCGAACAACTGGCC
GGCGATTGCGGCGTTCCACCAGGCGACGGAGCAGCATGGTCAGCTGATGACGCTTGCGGATCTGATCGGCTTCCACTTCGTGC
GCACGCTCGATGCCGGAATATATCCGCATGTCGAGC

Methylosinus

GACTTCTGGATTGACTGGAAGGATCGTCGTATGTGGCCGACGGTCGTGCCGATCCTTGGCGTGACCTTCTGCGCGGCGTCGCA
GGCTTTCTGGTGGGTTAACTTCCGCCTTCCGTTCGGCGCCGTTTTCGCGGCTCTTGGCCTCCTGATCGGCGAGTGGATCAACC
GCTACGTCAACTTCTGGGGCTGGACCTACTTCCCGATTTCGCTGGTGTTCCCGTCTGCTCTGATCGTTCCGGCGATCTGGCTC
GACGTGATCCTGCTGCTGTCGGGCTCCTATGTGATCACGGCGGTTGTCGGTTCGCTGGGCTGGGGTCTGCTGTTCTACCCGAA
CAACTGGCCGGCGATCGCCGCCTTCCACCAGGCGACCGAGCAGCATGGTCAGCTGATGACCCTGGCTGACTCTGATCGGCCTC
CACTACGTCCGCACGTCGATGCCGGAATACATCCGCATGTCGAGCGCGTA

Appendix 10

Sequence alignment with the program MUSCLE; in Phylip intermeaved format: (1:a phylogenetic tree)

```
20  570

10b/1-459                          GGGGACTGGGACTTCTGG--------------------------------
13d/1-460                          GGGGACTGGGACTTCTGGG-------------------------------
Methylocystis/1-451                ------------------GTTGACTGGAAGGATCGTCGTATGTGGCCGAC
16b/1-565                          GGGGACTGGGACTTCTGGGTCGACTGGAAGGATCGTCGTATGTGGCCGAC
15i/1-565                          GGTGACTGGGACTTCTGGGTTGACTGGAAGGATCGTCGTATGTGGCCGAC
15e/1-565                          GGAGACTGGGACTTCTGGGTTGACTGGAAGGATCGTCGTATGTGGCCGAC
16c/1-565                          GGGGACTGGGACTTCTGGGTTGACTGGAAGGATCGTCGTATGTGGCCGAC
Methylosinus/1-465                 ---------GACTTCTGGATTGACTGGAAGGATCGTCGTATGTGGCCGAC
4/1-479                            GGGGACTGGGACTTCTGG--------------------------------
12b/86-630                         TGGGGCGGGCGCTTCTGGGCGGACTGGAAAGACCGTCGTCTGTGGGTGAC
10c/1-565                          GGCGACTGGGACTTCTGGGTTGACTGGAAGGATCGTCGTTTGTGGGTGAC
15b/1-545                          GGAGACTGGGACTTCTGGGTTGACTGGAAAGACCGCCGTCTATGGGTGAC
1Methylomonasmethanica/1-495       ---GACTGGGACTTCTGGTCGGACTGGAAAGACCGTCGACTGTGGGTCAC
13a/1-545                          GGCGACTGGGACTTCTGGTCCGACTGGAAAGACCGTCGTCTGTGGGTCAC
16e/1-545                          GGGGACTGGGACTTCTGGACGGACTGGAAAGACAGACGTCTGTGGGTAAC
12c/91-655                         GGTGACTGGGACTTCTGGACTGACTGGAAAGACAGACGTCTGTGGGTAAC
1MAU31654Methylobacteralbus/1-495 ------------------ACCGATTGGAAAGACCGTCGTCTGTGGGTAAC
10a/1-545                          GGGGACTGGGACTTCTGGACCGACTGGAAAGACCGTCGTCTGTGGGTAAC
13b/1-545                          GGGGACTGGGACTTCTGGACTGACTGGAAAGACCGTCGTCTGTGGGTAAC
15d/1-545                          GGTGACTGGGACTTCTGGACTGACTGGAAAGACCGTCGTCTGTGGGTAAC

10b/1-459                          ------------------------CCTTCT----------------------
13d/1-460                          ------------------------CCTTCT----------------------
Methylocystis/1-451                GGTCGTGCCGATTCTCGGCGTGACCTTCTGCGCGGCGTCGCAGGCGTTCT
16b/1-565                          GGTCGTGCCGATTCTGGGCGTGACCTTCTGCGCGGCTTCGCAGGCTTTCT
15i/1-565                          GGTCGTGCCGATCCTGGGCGTGACCTTCTGCGCGGCGTCGCAGGCTTTCT
15e/1-565                          GGTCGTGCCGATCCTTGGCGTGACCTTCTGCGCGGCGTCGCAGGCTTTCT
16c/1-565                          GGTCGTGCCGATTCTTGGCGTGACCTTCTGCGCGGCGTCGCAGGCTTTCT
Methylosinus/1-465                 GGTCGTGCCGATCCTTGGCGTGACCTTCTGCGCGGCGTCGCAGGCTTTCT
4/1-479                            ------------------------CCTTCT----------------------
12b/86-630                         GGTGACGCCGATCGTGTTGATCACGTTTCCGGCGGCGGTACAGGCGTACT
10c/1-565                          GGTGGTGCCGATCGTGCTGGTAACTTTCCCGGCTGCGGTACAGGCCTTCT
15b/1-545                          CGTGGTGCCGATCGTGTTGGTCACCTTCCCGGCCGCCACCCAGGTGATGC
1Methylomonasmethanica/1-495       GGTGACCCCGATCGTACTGGTCACCTTCCCGGCGGCCGTACAATCCTACC
13a/1-545                          GGTGACCCCGATCGTGTTGGTCACCTTCCCGGCTGCCGTTCAATACTTCC
16e/1-545                          GGTTGTGCCGATCGTAGCTGTTACTTTCCCTGCGGCTGTGCAAGCTTTCG
12c/91-655                         CGTTTTACCGATTGTTGGTGTTACATTCCCTGCTGCCGTACAAGCCGTAC
1MAU31654Methylobacteralbus/1-495 CGTATTGCCAATCGTTGGTATTACTTTCCCAGCCGCTGTTCAAGCAGTTG
10a/1-545                          CGTATTGCCAATCGTTGGTATTACATTCCCGGCTGCGGTTCAAGCAGTTC
13b/1-545                          CGTATTGCCAATCGTTGGTATTACATTCCCAGCCGCTGTTCAAGCAGTTC
15d/1-545                          CGTATTGCCAATCGTTGGTATTACATTCCCAGCCGCTGTTCAAGCAGTTC

10b/1-459                          ----------TCTCCGCGTTCTGTTCGGCGCCGTTTTCGCGGCTCTGGGC
13d/1-460                          ----------TCTCCGCGTTCCGTTCGGCGCCGTTTTCGCGGCTCTGGGC
Methylocystis/1-451                GGTGGGTGAACTTCCGTCTGCCGTTTGGCGCGGTGTTCGCGGCTCTTGGC
16b/1-565                          GGTGGGTGAACTTCCGTCTGCCGTTCGGCGCGGTGTTTGCGGCTCTCGGC
15i/1-565                          GGTGGGTTAACTTCCGTCTTCCGTTCGGCGCCGTTTTCGCGGCTCTGGGC
15e/1-565                          GGTGGGTTAACTTCCGCCTTCCGTTCGGCGCCGTGTTCGCGGCTCTGGGC
```

```
16c/1-565                           GGTGGGTGAACTTCCGTCTGCCGTTCGGCGCGGTGTTCGCGGCTCTCGGC
Methylosinus/1-465                  GGTGGGTTAACTTCCGCCTTCCGTTCGGCGCCGTTTTCGCGGCTCTTGGC
4/1-479                             ----------TCTCCGCCTTCCGTTCGGCGCCGTTTTCGCGGCTCTGGGC
12b/86-630                          TGTGGGAGCGGTATCGTCAACCGTGGGGTGCGACGGTGTGCGTGCTGGGT
10c/1-565                           TGTGGGAGCGTTTCCGTCTGCCTTGGGGCGCGACGATTTGCGTACTGGGC
15b/1-545                           TGTGGGAGCGTCTGCGTCTTCCTTGGGGCGCGACGGTATGCGTATTGGCC
1Methylomonasmethanica/1-495       TGTGGGAGCGGTATCGTCTGCCCTGGGGAGCCACCGTGTGCGTCCTGGGT
13a/1-545                           TGTGGGAACGGTTCCGTCAGCCCTGGGGTGCGACCGTGTGCGTACTGGCG
16e/1-545                           CATGGGGTCGTTACCGTTTGCCTTGGGGCGCAACCGTTTGCGTTCTGGGC
12c/91-655                          TTTGGTATCGCTACCGTTTGCCGTTCGGTGCAGTTATTTCCGTTCTGGGT
1MAU31654Methylobacteralbus/1-495 TGTGGTATCGTTGGCGTCTGCCATTCGGCGCAATGCTGGCTGTTCTGGGC
10a/1-545                           TTTGGTACCGTTACCGTCTGCCATTCGGCGCTATGCTGGCTGTATTAGGT
13b/1-545                           TTTGGTACCGCTATCGTCTGCCATTCGGCGCAATGCTGGCTGTTCTTGGC
15d/1-545                           TTTGGTACGGCTATCGTCTGCCATTCGGCGCAATGCTGGCTGTTCTTGGC

10b/1-459                           CTCCTGATCGGCGAGTGGATCAATCGCTACGTCAACTTCTGGGGCTGGAC
13d/1-460                           CTTCTGATTGGCGAGTGGATTAACCGCTACGTCAACTTCTGGGGCTGGAC
Methylocystis/1-451                 CTTCTGATTGGCGAGTGGATCAACCGCTACGTCAACTTCTGGGGCTGGAC
16b/1-565                           CTTCTGATTGGCGAGTGGATCAACCGCTACGTCAACTTCTGGGGCTGGAC
15i/1-565                           CTGCTGATTGGCGAGTGGATCAACCGCTACGTCAACTTCTGGGGCTGGAC
15e/1-565                           CTCCTGATTGGCGAGTGGATCAACCGCTACGTCAACTTCTGGGGCTGGAC
16c/1-565                           CTGCTGATCGGCGAGTGGATCAACCGCTACGTCAACTTCTGGGGCTGGAC
Methylosinus/1-465                  CTCCTGATCGGCGAGTGGATCAACCGCTACGTCAACTTCTGGGGCTGGAC
4/1-479                             CTCCTGATTGGCGAGTGGATCAACCGCTACGTCAACTTCTGGGGCTGGAC
12b/86-630                          CTGTTGTTGGGTGAGTGGGTCAACCGCTACTTCAACTTCTGGGGCTGGAC
10c/1-565                           CTGCTGTTTGGTGAGTGGGTCAACCGCTACTTCAACTTCTGGGGCTGGAC
15b/1-545                           CTGTTGTTTGGTGAGTGGATCAACCGCTACTTCAACTTCTGGGGTTGGAC
1Methylomonasmethanica/1-495       CTGCTGCTGGGCGAGTGGATCAACCGTTATTTCAACTTCTGGGGCTGGAC
13a/1-545                           CTGCTCTTCGGTGAATGGGTCAACCGCTACTTTAACTTCTGGGGCTGGAC
16e/1-545                           TTGCTGTTCGGTGAATGGGTTAACAGATACTTCAACTTCTGGGGTTGGAC
12c/91-655                          CTGCTGTTTGGTGAGTGGGTTAACAGATACTTCAATTTCTGGGGATGGAC
1MAU31654Methylobacteralbus/1-495 CTGCTGTTCGGCGAATGGGTTAACAGATATTTCAACTTCTGGGGATGGAC
10a/1-545                           CTGCTGTTCGGCGAATGGGTCAACAGATACTTCAACTTCTGGGGATGGAC
13b/1-545                           CTGTTGTTCGGTGAATGGGTTAACAGATATTTCAACTTCTGGGGATGGAC
15d/1-545                           CTGTTGTTCGGTGAATGGGTTAACAGATATTTCAACTTCTGGGGATGGAC

10b/1-459                           CTATTTCCCGATCAGCCTCGTGTTCCCGTCCGCTCTGATCGTCCCG---G
13d/1-460                           GTATTTCCCGATCAGCCTCGTTTTCCCGTCTGCCCTGATCGTTCCG---G
Methylocystis/1-451                 CTACTTCCCGATCAGCCTTGTGTTCCCGTCTGCGTTGATCGTTCCG---G
16b/1-565                           CTACTTCCCGATCAGCCTCGTGTTCCCGTCTGCTCTGATCGTTCCG---G
15i/1-565                           GTATTTCCCGATCAGCCTCGTGTCCCCGTCCGCTCTGATCGTTCCG---G
15e/1-565                           GTACTTCCCGATCAGCCTCGTGTTCCCGTCCGCTCTGATCGTTCCG---G
16c/1-565                           CTACTTCCCGATCAGCCTGGTGTTCCCGTCTGCTCTGATCGTTCCG---G
Methylosinus/1-465                  CTACTTCCCGATTTCGCTGGTGTTCCCGTCTGCTCTGATCGTTCCG---G
4/1-479                             CTATTTCCCGATCAGCCTCGTGTTCCCGTCCGCTCTGATCGTTCCG---G
12b/86-630                          GTATTTTCCGGTGGACTTCGTATTTCCGG---CGATTTTGACGCCGGGTG
10c/1-565                           CTACTTCCCGATCACCCTGTGCTTCCCGT---CGCAGATCGTACCGGGCG
15b/1-545                           CTACTTCCCGATCACCCTGTGCTTCCCCT---CGCAGATCGTGCCGGGCG
1Methylomonasmethanica/1-495       CTACTTCCCGATCAACTTCGTGTTCCCTG---CCTCGCTGGTGCCGGGCG
13a/1-545                           CTACTTCCCCGTGAACTTCGTGTTCCCGG---CGGTCCTGGTTCCGGGCG
16e/1-545                           ATACTTCCCGGTTAACTTCGTATTCCCAT---CACAATTGATTCCTAGCG
12c/91-655                          TTATTTCCCAGTGAATTTCGTATTCCCAT---CACAATTCGTGCCAGGCG
1MAU31654Methylobacteralbus/1-495 TTACTTCCCAGTTAACTTCGTGTTCCCAT---CACAATTCGCTCCAGGCG
10a/1-545                           TTACTTCCCGGTTAACTTTGTATTCCCAT---CACAATTCGTTCCAGGCG
13b/1-545                           TTACTTCCCAGTTAACTTTGTATTCCCAT---CACAATTCGTTCCAGGCG
15d/1-545                           TTACTTCCCAGTTAACTTTGTATTCCCAT---CACAATTCGTTCCAGGCG
```

```
10b/1-459                           CGATCTGGCTCGACGTGATCCTGCTTCTGTCGGGTTCCTATGTGATCACG
13d/1-460                           CGATCTGGCTTGACGTGATCCTGCTCCTGTCGGGTTCCTATGTGATCACG
Methylocystis/1-451                 CGATCTGGCTTGACGTGATCCTGCTTCTGTCGGGGTCCTATGTGATCACG
16b/1-565                           CGATCTGGCTCGACGTGATCCTGCTTCTGTCGGGCTCCTATGTGATCACG
15i/1-565                           CGATCTGGCTCGACGTGATCCTGCTGCTGTCGGGCTCCTATGTGATCACG
15e/1-565                           CGATCTGGCTCGACGTGATCCTGCTTCTGTCGGGTTCCTATGTGATCACG
16c/1-565                           CGATCTGGCTGGACGTGATCCTGCTGCTGTCGGGCTCCTATGTGATCACG
Methylosinus/1-465                  CGATCTGGCTCGACGTGATCCTGCTGCTGTCGGGCTCCTATGTGATCACG
4/1-479                             CGATCTGGCTTGACGTGATCCTGCTTCTGTCGGGTTCCTATGTGATCACG
12b/86-630                          CGATTTTGCTGGACACGATGCTGTTGCTGAGCGGCAGCTATCTGTTCACG
10c/1-565                           CCATCCTGCTCGACACGGCTCTGATGCTGAGCACCAGCTACCTGTTCACC
15b/1-545                           CCATCCTGCTCGACGTCGTTCTGTTGCTCTCCGGCAGCTACCTGCTCACC
1Methylomonasmethanica/1-495        CCATCATCCTGGACACCGTGCTGATGCTGTCGGGCAGCTACCTGTTCACC
13a/1-545                           CGATCCTGCTCGACGTTATCCTGATGCTCTCGGGCAGCTACCTGTTTGCT
16e/1-545                           CTATCCTGCTGGACGTGGTTCTGTTGCTGTCCAACAGCTACACCTTCACC
12c/91-655                          CAATCGTTCTTGGTGTTATCCTGATGCTGTCCAACAGCATGCAATTAACA
1MAU31654Methylobacteralbus/1-495   CAATCGTTCTGGACGTAATTCTGATGCTGTCTAACAGCATGCAGTTGACT
10a/1-545                           CACTCGTTCTGGACGTGATCCTGATGTTGTCAAACAGCATGCAGTTGACT
13b/1-545                           CACTGGTTCTGGACGTAATTCTGATGCTGTCTAACAGCATGCAGTTGACT
15d/1-545                           CACTGGTTCTGGACGTAATTCTGATGCTGTCTAACAGCATGCAGTTGACT

10b/1-459                           GCGGTTGTCGGTTCGCTGGGCTGGGGTCTGCTGTTCTACCCGAACAACTG
13d/1-460                           GCGGTTGTCGGTTCGCTGGGCTGGGGTCTGCTGTTCTACCCGAACAACTG
Methylocystis/1-451                 GCGATTGTTGGTTCGCTGGGCTGGGGTCTGTTGTTCTACCCGAACAACTG
16b/1-565                           GCGGTTGTCGGCTCGCTGGGCTGGGGTCTGCTGTTCTATCCGAACAACTG
15i/1-565                           GCGGTTGTCGGTTCGCTGGGCTGGGGTCTGCTGTTCTATCCGAACAGCTG
15e/1-565                           GCGGTTGTCGGTTCGCTGGGCTGGGGTCTGCTGTTCTACCCGAACAACTG
16c/1-565                           GCGGTTGTCGGTTCGCTGGGCTGGGGTCTGCTGTTCTATCCGAACAACTG
Methylosinus/1-465                  GCGGTTGTCGGTTCGCTGGGCTGGGGTCTGCTGTTCTACCCGAACAACTG
4/1-479                             GCGGTTGTCGGTTCGCTGGGCTGGGGTCTGCTGTTCTACCCGAACAACTG
12b/86-630                          GCGATCATTGGCGGTTTGGGCTGGGGTTTGATTTTCTACCCGGGCAACTG
10c/1-565                           GCGATCGTCGGTGCGATGGCTTGGGGCTTGGTTTTCTATCCGGGCAACTG
15b/1-545                           GCCATCGTGGGCGGCATGGCTTGGGGCTTGATCTTCTACCCGGGCAACTG
1Methylomonasmethanica/1-495        GCGATCGTCGGTGCGATGGGCTGGGGTCTGACCTTCTACCCGGGCAACTG
13a/1-545                           GCCATCATCGGCGGTCTGGCTTGGGGCCTGATTTTCTATCCGGGCAACTG
16e/1-545                           GCTGTTGCTGGCGCTATGGGTTGGGGCTTGATTTTCTATCCTAGCAACTG
12c/91-655                          GCCGTTATCGGTGGTTTGGGCTATGGTCTGTTGTTCTATCCTGGCAACTG
1MAU31654Methylobacteralbus/1-495   GCGGTTCTGGGCGGCTTGGCTTATGGGCTGTTGTTCTATCCTGGCAACTG
10a/1-545                           GCTGTGCTGGGCGGCTTGGCTTATGGTCTGTTGTTCTATCCTGGCAACTG
13b/1-545                           GCTGTTATCGGCGGTCTGGCATACGGCTTGTTGTTCTATCCTGGCAACTG
15d/1-545                           GCTGTTATCGGCGGTCTGGCATACGGCTTGTTGTTCTATCCTGGCAACTG

10b/1-459                           GCCGGCGATTGCGGCGTTCCACCAGGCGACCGAGCAGCATGGTCAGCTGA
13d/1-460                           GCCGGCGATCGCCGCCTTCCACCAGGCGACCGAGCAGCATGGTCAGCTGA
Methylocystis/1-451                 GCCGGCGATTGCGGCGTTCCACCAGGCGACGGAGCAGCATGGTCAGCTGA
16b/1-565                           GCCGGCGATCGCCGCCTTCCACCAGGCGACCGAGCAGCATGGTCAGCTGA
15i/1-565                           GCCGGCGATTGCGGCGTTCCACCAGGCGACTGAGCAGCACGGCCAGCTGA
15e/1-565                           GCCGGCGATTGCGGCGTTCCACCAGGCGACCGAGCAGCATGGTCAGCTGA
16c/1-565                           GCCGGCGATCGCCGCCTTCCACCAGGCGACCGAGCAGCATGGTCAGCTGA
Methylosinus/1-465                  GCCGGCGATCGCCGCCTTCCACCAGGCGACCGAGCAGCATGGTCAGCTGA
4/1-479                             GCCGGCGATCGCCGCCTTCCACCAGGCGACCGAGCAGCATGGTCAGCTGA
12b/86-630                          GCCGATCATTGCGCCGCTGCACGTACCGGTGGAGTACAGCGGGATGCTGA
10c/1-565                           GCCGGTGATTGCACCGTACCACGTGCCGGTAGAGTACAACGGCATGCTGA
15b/1-545                           GCCGGTGATCGCGCCGTACCACGTGCCGGTGGAATACAACGGCATGCTGC
1Methylomonasmethanica/1-495        GCCGATCATCGCGCCGCTGCACGTGCCGGTGGAATACAACGGCATGCTGA
13a/1-545                           GCCGGTGATTGCTCCGCTGCACGTGCCGGTGGAATACAACGGCATGTTGA
16e/1-545                           GCCGGTTATTGGTCCATTACACGTGCCTGTTGAATATAACGGCATGATGA
```

```
12c/91-655                          GCCTGTCATCGCTCCATTGCACGTGCCTGTTGAATATAACGGTATGGTTA
1MAU31654Methylobacteralbus/1-495 GCCGGTCATCGCTCCATTGCACGTGCCAGTTGAATACAACGGGATGGTAA
10a/1-545                           GCCTGTCATCGCTCCATTGCACGTGCCAGTTGAATACGACGGCATGGTAA
13b/1-545                           GCCTGTCATCGCTCCATTGCACGTGCCTGTTGAATACAACGGCATGGTAA
15d/1-545                           GCCTGTCATCGCTCCATTGCACGTGCCTGTTGAATACAACGGCATGGTAA

10b/1-459                           TGACGCTTGCTGAC-CTGATCGGCCTCCACTACGTCCGCACG-TCGATGC
13d/1-460                           TGACCCTCGCCGAC-CTGATTGGTCTGCACTACGTCCGCACG-TCGATGC
Methylocystis/1-451                 TGACGCTTGCGGAT-CTGATCGGCTTCCACTTCGTGCGCACGCTCGATGC
16b/1-565                           TGACGCTGGCCGAC-CTCATCGGCTTCCACTTCGTCCGCACC-TCGATGC
15i/1-565                           TGACGCTTGCTGAC-CTGATCGGCCTCCACTACGTCCGCACG-TCGATGC
15e/1-565                           TGACCCTGGCTGAC-CTGATCGGCCTTCACTGTGTCCGCACG-TCGAAGC
16c/1-565                           TGACGCTTGCTGAC-CTGATCGGCCTCCACTACGTCCGCACG-TCGATGC
Methylosinus/1-465                  TGACCCTGGCTGACTCTGATCGGCCTCCACTACGTCCGCACG-TCGATGC
4/1-479                             TGACCCTGGCTGAC-CTGATCGGCCTCCACTACGTCCGCACG-TCGATGC
12b/86-630                          TGTCGATTGCCGAC-ATTCAGGGTTACAACTACGTGCGTACC-GGAACCC
10c/1-565                           TGTCGGTTGCCGAC-CTGCTGGGTTACCACTATGTCCGTACG-GGTACGC
15b/1-545                           TGTCGGTGGCCGAC-TTGCTGGGCTACAACTATGTTCGTACC-GGCACCC
1Methylomonasmethanica/1-495        TGTCGATCGCCGAC-ATCCAGGGTTACAACTATGTGCGTACG-GGTACGC
13a/1-545                           TGTCCATTGCCGAC-ATCCAGGGCTACAACTATGTTCGTACC-GGTACCC
16e/1-545                           TGACTTTGGCTGAC-TTACAAGGTTACCACTATGTAAGAACC-GGTACTC
12c/91-655                          TGACGCTGGCTGAC-TTGCAAGGTTACCACTATGTAAGAACC-GGTACCC
1MAU31654Methylobacteralbus/1-495 TGACCCTGGCTGAC-TTGCAAGGTTACCACTATGTAAGAACC-GGTACTC
10a/1-545                           TGACCCTGGCTGAC-TTGCAAGGTTACCACTACGTAAGAACT-GGTACTC
13b/1-545                           TGACGCTGGCTGAC-TTGCAAGGTTACCACTATGTAAGAACT-GGTACTC
15d/1-545                           TGACGCTGGCTGAC-TTGCAAGGTTACCACTATGTAAGAACT-GGTACTC

10b/1-459                           CGGAATACATCCGCATGGTCGAGCGCGGCACGCTGCGCACGTTCGGTAAG
13d/1-460                           CGGAATACATCCGCATGGTCGAGCGCGGCACGCTGCGCACGTTCGGTAAA
Methylocystis/1-451                 CGGAATATATCCGCATG-TCGAGC--------------------------
16b/1-565                           CGGAATACATCCGCATGGTCGAGCGCGGCACGCTGCGCACCTTCGGTAAA
15i/1-565                           CGGAATACATCCGCATGGTCGAGCGCGGCACGCTGCGCACGTTCGGTAAA
15e/1-565                           CGGAATACATCCGCATGGTCGAGCGCGGCACGCTGCGCACGTTCGGTAAA
16c/1-565                           CGGAATACATCCGCATGGTCGAGCGCGGCACGCTGCGCACGTTCGGTAAA
Methylosinus/1-465                  CGGAATACATCCGCATG-TCGAGC----------GCGTA-----------
4/1-479                             CGGAATACATCCGCATGGTCGAGCGCGGCACGCTGCGCACGTTCGGTAAG
12b/86-630                          CGGAATACATCCGGATGGTCGAGAAGGGCACGCTGCGTACGTTTGGTAAA
10c/1-565                           CTGAGTACATCCGTATGGTCGAGAAGGGCACCCTGCGTACCTTCGGTAAA
15b/1-545                           CCGAGTACATCCGCATGGTTGGGAAAGGCACCCTGCGTACCTTCGGTAAA
1Methylomonasmethanica/1-495        CTGAGTACATCCGCATGGTAGAGAAGGGCACCCTGCGTACCTTCGGTAAG
13a/1-545                           CCGAGTACATCCGCATGGTTGAGAAAGGCACCCTGCGTACCTTCGGTAAA
16e/1-545                           CTGAATACATCCGTATGGTTGAAAAAGGTACATTGAGAACTTTCGGTAAA
12c/91-655                          CAGAGTACATCAGAATGGTTGAAAAAGGTACGTTGAGAACATTTGGTAAA
1MAU31654Methylobacteralbus/1-495 CAGAATATATCCGGATGGTTGAAAAAGGTACTCTGAGAACTTTCGGTAAG
10a/1-545                           CAGAGTACATCAGAATGGTTGAAAAAGGTACTTTGAGAACATTCGGTAAA
13b/1-545                           CAGAATACATCAGAATGGTTGAAAAAGGTACTTTGAGAACTTTCGGTAAG
15d/1-545                           CAGAATACATCAGAATGGTTGAAAAAGGTACTTTGAGAACTTTCGGTAAG

10b/1-459                           GACGTTGCTCCGGAAGGGCGAAT----------TCCAGCACACTGGCGGC
13d/1-460                           GACGTTGCTCCGGAAGGGCGAAT----------TCCAGCACACTGGCGGC
Methylocystis/1-451                 --------------------------------------------------
16b/1-565                           GACGTTGCGCCGGAAGGGCGAATTCTGCAGATATCCATCACACTGGCGGC
15i/1-565                           GACGTTGCTCCGGAAGGGCGAATTCTGCAGATATCCATCACACTGGCGGC
15e/1-565                           GACGTTGCTCCGGAAGGGCGAATTCTGCAGATATCCATCACACTGGCGGC
16c/1-565                           GACGTTGCGCCGGAAGGGCGAATTCTGCAGATATCCATCACACTGGCGGC
Methylosinus/1-465                  --------------------------------------------------
4/1-479                             GACGTTGCTCCGGAAGGGCGAATTCTGCAGATATCCATCACACTGGCGGC
```

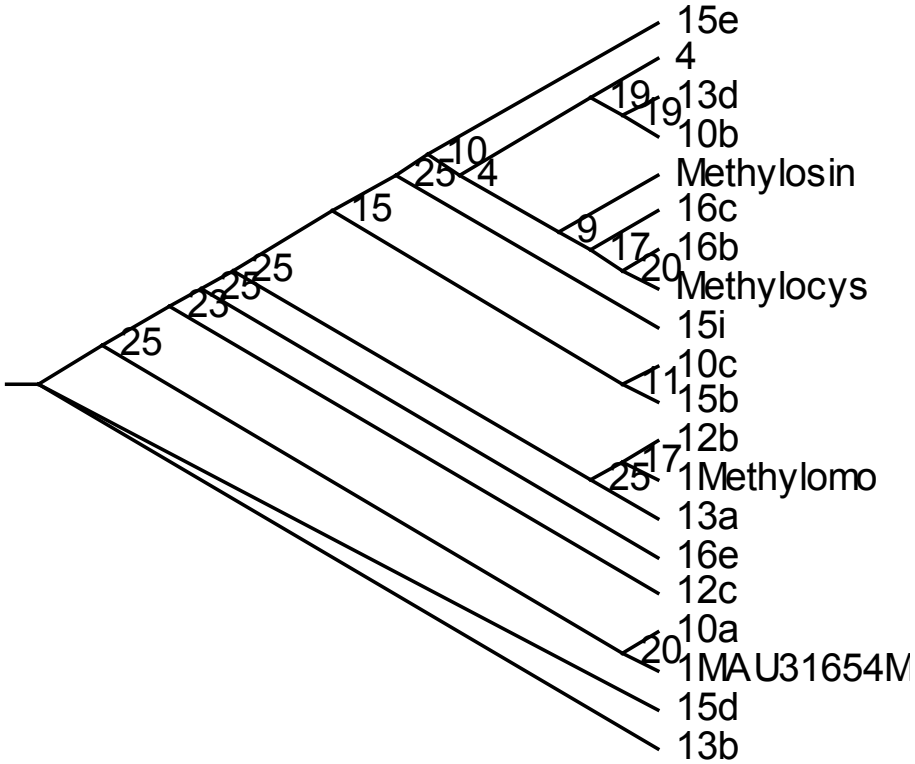
```
12b/86-630                          GACGTTGCGCCGGAAGGGCGAAT----------TCCAGCACACTGGCGGC
10c/1-565                           GACGTTGCTCCGGAAGGGCGAATTCTGCAGATATCCATCACACTGGCGGC
15b/1-545                           GACGTTGCGCCGGAAGGGCGAAT----------TCCAGCACACTGGCGGC
1Methylomonasmethanica/1-495        GAC-------------------------------------------
13a/1-545                           GACGTTGCTCCGGAAGGGCGAAT----------TCCAGCACACTGGCGGC
16e/1-545                           GACGTTGCTCCGGAAGGGCGAAT----------TCCAGCACACTGGCGGC
12c/91-655                          GACGTTGCGCCGGAAGGGCGAATTCTGCAGATATCCATCACACTGGCGGC
1MAU31654Methylobacteralbus/1-495 GACGTTGCTCCAGTA------------------TCC--------------
10a/1-545                           GACGTTGCTCCGGAAGGGCGAAT----------TCCAGCACACTGGCGGC
13b/1-545                           GACGTTGCTCCGGAAGGGCGAAT----------TCCAGCACACTGGCGGC
15d/1-545                           GACGTTGCGCCGGAAGGGCGAAT----------TCCAGCACACTGGCGGC

10b/1-459                           CGTTA----------CTAGT
13d/1-460                           CGTTA----------CTAGT
Methylocystis/1-451                 --------------------
16b/1-565                           CGCTCGAGCATGCATCTAGA
15i/1-565                           CGCTCGAGCATGCATCTAGA
15e/1-565                           CGCTCGAGCATGCATCTAGA
16c/1-565                           CGCTCGAGCATGCATCTAGA
Methylosinus/1-465                  --------------------
4/1-479                             CGCTCGAGCATGCATCTAGA
12b/86-630                          CGTTA----------CTAGT
10c/1-565                           CGCTCGAGCATGCATCTAGA
15b/1-545                           CGTTA----------CTAGT
1Methylomonasmethanica/1-495        --------------------
13a/1-545                           CGTTA----------CTAGT
16e/1-545                           CGTTA----------CTAGT
12c/91-655                          CGCTCGAGCATGCATCTAGA
1MAU31654Methylobacteralbus/1-495 --------------------
10a/1-545                           CGTTA----------CTAGT
13b/1-545                           CGTTA----------CTAGT
15d/1-545                           CGTTA----------CTAGT
```

Appendix 11

Phylogenetic tree in Newick format

(((((((15e:0.012941,((4:0.000000,(13d:0.031721,10b:0.024059)19:0.014699)19:0.012852,(Methylosin:0.019377,(16c:0.008894,(16b:0.018828,Methylocys:0.051683)20:0.008716)17:0.022133)9:0.007158)4:0.008882)10:0.021596,15i:0.007799)25:0.599955,(10c:0.084924,15b:0.110246)11:0.024747)15:0.040353,((12b:0.193643,1Methylomo:0.095768)17:0.043659,13a:0.087311)25:0.057974)25:0.212213,16e:0.074704)25:0.146622,12c:0.106425)23:0.049976,(10a:0.045890,1MAU31654M:0.051379)20:0.014933)25:0.019381,15d:0.005597,13b:0.000000);

Slanted cladogram with bootstrap values



Appendix 12

Type I methanotrophs (pmoA gene)

Methylomicrobium

GGGGACTGGGACTTCTGGACCGATTGGAAAGACCGTCGTCTGTGGGTAACCGTATTGCCAATCGTTGGTA
TTACTTTCCCAGCCGCTGTTCAAGCAGTTGTGTGGTATCGTTGGCGTCTGCCATTCGGCGCAATGCTGGC
TGTTCTGGGCCTGCTGTTCGGCGAATGGGTTAACAGATATTTCAACTTCTGGGGATGGACTTACTTCCCA
GTTAACTTCGTGTTCCCATCACAATTCGTTCCAGGCGCAATCGTTCTGGACGTAATTCTGATGCTGTCTA
ACAGCATGCAGTTGACTGCGGTTCTGGGCGGCTTGGCTTATGGCCTGTTGTTCTATCCTGGCAACTGGCC
GGTCATCGCTCCATTGCACGTGCCAGTTGAATACAACGGCATGGTAATGACCCTGGCTGACTTGCAAGGT
TACCACTATGTAAGAACCGGTACTCCAGAATATATCCGGATGGTTGAAAAAGGTACTCTGAGAACTTTCG
GTAAGGACGTTGCTCCAGTATCTGCCTTCTTCTCCGCCTTC

Methylocaldum

TCCGACTGGAAAGACCGTCGTCTGTGGGTCACGGTCACCCCGATCGTGTTGGTGACGTTCCCGGCGGCGA
TCCAGGCGTGGACCTGGGAACGGTTCCGCAACCCGTGGGGCGCAACCATGGCCGTTCTGGCGCTGCTCTT
CGGCGAATGGGTCAACCGCTACTTCAACTTCTGGGGCTGGACCTACTTCCCGATCAACTTCGTATTCCCG
GCCATTCTGGTTCCGGGTGCGATCTTGCTCGACACCTTCCTGATGCTTTCGGGCAGCTACCTGTTCACGG
CGATCGTGGGCGGCATGGCTTGGGGCCTGATTTTCTATCCGGGCAACTGGCCGATGATCGCACCGCTGCA
CGTGCCGGTGGAATACAACGGCATGCTCATGTCGATTGCCGACTTGCAGGGCTACCACTATGTCCGTACC
GGTACCCCGCGAGTACATCCCGC

Methylothermus

TTGTGGGTGACGGTGTACCCGATTGTGATGATCACCTTCCCGGCGGCGGTGCAAGCGGTGCTGTGGGAGC
GTTTGCGGCTGCCGTGGGGGGCGACGGTCTGTGTTTTGGGCATCCTGTTTGGGGAATGGGTGAACCGTTA
CTTTAACTTCTGGATGTGGACCTATTTCCCCATTAACTTTGTTTTCCCGACGGCGGCGGTGCACATGGCG
ATCTTTTTAGACGTGGTGCTGATGCTGTCTGGGAGTTTCTTGTTCACGGCGGTGATTGGCGGGTTAGGCT
GGGGTTTGTTGATGTATCCTGGCAACTGGCCGATTATTGCGCCTTTGCACGTGCCGGTGGAATACAACGG
CATGCTGATGTCGATTGCCGACATTCAAGGCTACCACTACGTGCGCACGGGGACGCCGGAATACATTCGG
ATGGTAGAGAAAGGCACGCTGCGTACCTTTGGTAAGGACA

Methylosarcina

TGGGGGAACTTCTGGGGGTGGACTTATTTCCCAGTAAACTTCGTGTTCCCTTCAAACTTCATGCCTGGCG
CTATCGTTCTGGACGTTATCCTGATGCTGTCCAACAGCATGCAGCTGACGGCTGTTATCGGTGGTTTAGG
CTACGGCTTGTTGTTCTATCCTGGTAACTGGCCAGTCATCGCTCCATTGCACGTCCCTGTTGAATACAAT
GGTATGGTAATGACTCTGGCTGACTTGCAAGGTTACCACTATGTAAGAACTGGTACACCTGAATACATCC
GGATGGTTGAAAAAGGTACACTGAGAACTTTCGGTAACGACGTCGCCCCCAG

Methylohalobius

TCCGACTGGAAAGACCGTCGTCTGTGGGTGACCGTCTACCCGATCGTCATGATCACCTTCCCGGCCGCGG
TGCAAGCGGTGATTTGGGAACGTCTGCGCCTGCCTTTCGGGGCCACCATTTCCATTCTGGGCATTCTGCT
GGGTGAATGGATCAACCGCTACTTCAACTTCTGGGGATGGACCTACTTCCCGATCAACTTCGTCTTCCCG
ACGGCGGCGGTCCACATGGCCATCTTCCTGGACGTGGTCCTGATGCTGTCCAGCAGCTTCCTGTTCACGG
CCGTGGTTGGCGGCTTGGGCTGGGGTCTGTTGATGTACCCGGGCAACTGGCCGGTGATCGCACCGCTGCA
CGTGCCTGTGGAATACAACGGCATGCTGATGTCCGTGGCCGACATCCAAGGCTACCATTACGTCCGTACG
GGTACGCCCGAATACATCCGGATGGTTGAAAAAGGCACCCTGCGTACCTTCGGTAAAGACGTGGCGCCGG
TATCC

Methylosoma

GACTGGGACTTCTGGACTGACTGGAAAGACAGACGTCTTTGGGTAACTGTTGCCCCGATCGTTTCTATTA
CTTTCCCTGCTGCTGTTCAAGCTTGCTTGTGGTGGCGTTACCGTTTGCCATTCGGCGCTGTTGTTTGTAT
CTTAGGCTTGCTTTTGGGTGAGTGGATCAACAGATACTTGAACTTCTGGGGCTGGACTTACTTCCCTGTT
AACTTCGTGTTCCCATCAAACTTGATGCCAGGCGCTATCATTCTTGATGTTGTATTGTTGTTGTCAAACA

GCATGGTATTGACTGCCGTTATCGGTGGTTTGGCTTATGGTTTGGTTTTCTACCCAGGCAACTGGCCAAT
CATTGCTCCATTGCACGTTCCTGTTGAATACAATGGCATGATGATGACAATTGCTGACTTACAAGGTTAC
CACTACGTAAGAACTGGTACACCTGAGTACATCAGAATGGTTGAAAAAGGTACCTTAAGAACATTCGGTA
AAGACGTTGCTCCAGTATCAGCCTTCTTCTCAGCCTTC

Methylococcus

TCGGATTGGAAAGATCGTCGACTGTGGGTCACGGTGACCCCGATCGTGTTGATCACGTTCCCGGCGGCGG
TACAGTCGTACCTGTGGGAGCGGTTTCGTCAACCGTGGGGTGCTACGGTCTGCGTACTGGGCCTGTTGAT
GGGTGAGTGGATCAACCGTTACTTCAACTTCTGGGGCTGGACCTACTTCCCGGTCAATTTCGTGTTCCCG
GCGATTCTGACCCCGGGCGCGATCCTGCTCGACACGATGTTGATGCTGAGCGGCAGCTACCTGTTCACGG
CGATCGTTGGCGCGATGGGTTGGGGCCTGATTTTCTATCCGGGCAACTGGCCGATCATTGCACCGATTCA
CGTGCCGGTGGAATACAGCGGCATGCTGATGTCGATTGCCGACATTCAGGGTTACAACTACGTGCGTACC
TGGTACGCCGGAATACATCCGCAGCTGAGAA

Type II methanotrophs (pMOA gene)

Methanocapsa

ATCGACTGGAAAGATCGCCGCTTCTGGCCGACGGTTCTTCCGATCGTGCTCGTCACGTTCCCGGCCGCCG
CTCAGGCCTATTTCTGGGAAAGCTTCCGCCTTCCTTTCGGCGCGACCTTCCTGGTCCTCGGCCTTCTCTT
CGGTGAATGGGTCAACCGCTACACCAATTTCTGGGGTTGGACCTATTTCCCGATCAGCCTCGTTTGGCCG
ACCTCTTTGGTCCCGGCTGCTCTGTTCCTCGATATCGTGCTCTTGCTGAGCCGCAGCTTCATCGTCACCG
CGATCGTCGGCGCGATGGGCTGGGGTCTTCTCCTCTATCCCTCCAACTGGCCGATTCTTGCGCCTTATCA
TCAGGCGACGGAACAATATGGCCTTCTCATGTCTCTCGCCGACTTGATTGGCTTTGAATACGTCCGCACC
TCGATGCCTGAATATCTCCGCATCGTCGAGCGCGGCACGATGCGCACCTTCGGTAAGGACGTCGTCGCGG
TTGCGGCCTTCTTCTCAGCC

Appendix 13

Sequence alignment from the program MUSCLE, in Phylip interleaved format:
(2:a phylogenetic tree)

```
28 800

Methanocap ----------------------------------------------------
10b        ----------------------------------------------------
13d        ----------------------------------------------------
Methylocys ----------------------------------------------------
16b        ----------------------------------------------------
15i        ----------------------------------------------------
15e        ----------------------------------------------------
16c        ----------------------------------------------------
Methylosin ----------------------------------------------------
4          ----------------------------------------------------
16e        ----------------------------------------------------
Methylosom ----------------------------------------------------
12c        TGGAAACAAAGNTTGACCATGATTACGCCAAGCTTGGTACCGAGCTCGGA
Methylosar ----------------------------------------------------
Methylobac ----------------------------------------------------
Methylomic ----------------------------------------------------
10a        ----------------------------------------------------
13b        ----------------------------------------------------
15d        ----------------------------------------------------
10c        ----------------------------------------------------
15b        ----------------------------------------------------
13a        ----------------------------------------------------
Methylocal ----------------------------------------------------
Methylomon ----------------------------------------------------
12b        -----NNNNNTCGNNNATNCTGNTTNNTGGGCCCTCTAGATGCATGCTCG
Methylococ ----------------------------------------------------
Methylothe ----------------------------------------------------
Methylohal ----------------------------------------------------

           ----------------------------------------------------
           ----------------------------------------GGGGACTGGG
           ----------------------------------------GGGGACTGGG
           ----------------------------------------------------
           ----------------------------------------GGGGACTGGG
           ----------------------------------------GGTGACTGGG
           ----------------------------------------GGAGACTGGG
           ----------------------------------------GGGGACTGGG
           ----------------------------------------------------G
           ----------------------------------------GGGGACTGGG
           ----------------------------------------GGGGACTGGG
           -------------------------------------------GACTGGG
           TCCACTAGTAACGGCCGCCAGTGTGCTGGAATTCGCCCTTGGTGACTGGG
           ----------------------------------------------------
           ----------------------------------------------------
           ----------------------------------------GGGGACTGGG
           ----------------------------------------GGGGACTGGG
           ----------------------------------------GGGGACTGGG
           ----------------------------------------GGTGACTGGG
           ----------------------------------------GGCGACTGGG
```

```
----------------------------------------GGAGACTGGG
----------------------------------------GGCGACTGGG
--------------------------------------------------
-----------------------------------------GACTGGG
AGCGGCCGCCAGTGTGATGGATATCTGCAGAATTCCCCCTTGGGGCGGGC
--------------------------------------------------
--------------------------------------------------
--------------------------------------------------

--------ATCGACTGGAAAGATCGCCGCTTCTGGCCGACGGTTCTTCCG
ACTTCTGG----------------------------------------
ACTTCTGGG---------------------------------------
--------GTTGACTGGAAGGATCGTCGTATGTGGCCGACGGTCGTGCCG
ACTTCTGGGTCGACTGGAAGGATCGTCGTATGTGGCCGACGGTCGTGCCG
ACTTCTGGGTTGACTGGAAGGATCGTCGTATGTGGCCGACGGTCGTGCCG
ACTTCTGGGTTGACTGGAAGGATCGTCGTATGTGGCCGACGGTCGTGCCG
ACTTCTGGGTTGACTGGAAGGATCGTCGTATGTGGCCGACGGTCGTGCCG
ACTTCTGGATTGACTGGAAGGATCGTCGTATGTGGCCGACGGTCGTGCCG
ACTTCTGG----------------------------------------
ACTTCTGGACGGACTGGAAAGACAGACGTCTGTGGGTAACGGTTGTGCCG
ACTTCTGGACTGACTGGAAAGACAGACGTCTTTGGGTAACTGTTGCCCCG
ACTTCTGGACTGACTGGAAAGACAGACGTCTGTGGGTAACCGTTTTACCG
--------------------------------------------------
--------ACCGATTGGAAAGACCGTCGTCTGTGGGTAACCGTATTGCCA
ACTTCTGGACCGATTGGAAAGACCGTCGTCTGTGGGTAACCGTATTGCCA
ACTTCTGGACCGACTGGAAAGACCGTCGTCTGTGGGTAACCGTATTGCCA
ACTTCTGGACTGACTGGAAAGACCGTCGTCTGTGGGTAACCGTATTGCCA
ACTTCTGGACTGACTGGAAAGACCGTCGTCTGTGGGTAACCGTATTGCCA
ACTTCTGGGTTGACTGGAAGGATCGTCGTTTGTGGGTGACGGTGGTGCCG
ACTTCTGGGTTGACTGGAAAGACCGCCGTCTATGGGTGACCGTGGTGCCG
ACTTCTGGTCCGACTGGAAAGACCGTCGTCTGTGGGTCACGGTGACCCCG
--------TCCGACTGGAAAGACCGTCGTCTGTGGGTCACGGTCACCCCG
ACTTCTGGTCGGACTGGAAAGACCGTCGACTGTGGGTCACGGTGACCCCG
GCTTCTGGGCGGACTGGAAAGACCGTCGTCTGTGGGTGACGGTGACGCCG
--------TCGGATTGGAAAGATCGTCGACTGTGGGTCACGGTGACCCCG
-----------------------------TTGTGGGTGACGGTGTACCCG
--------TCCGACTGGAAAGACCGTCGTCTGTGGGTGACCGTCTACCCG

ATCGTGCTCGTCACGTTCCCGGCCGCCGCTCAGGCCTATTTCTGGGAAAG
-------------CCTTCT-------------------------------
-------------CCTTCT-------------------------------
ATTCTCGGCGTGACCTTCTGCGCGGCGTCGCAGGCGTTCTGGTGGGTGAA
ATTCTGGGCGTGACCTTCTGCGCGGCTTCGCAGGCTTTCTGGTGGGTGAA
ATCCTGGGCGTGACCTTCTGCGCGGCGTCGCAGGCTTTCTGGTGGGTTAA
ATCCTTGGCGTGACCTTCTGCGCGGCGTCGCAGGCTTTCTGGTGGGTTAA
ATTCTTGGCGTGACCTTCTGCGCGGCGTCGCAGGCTTTCTGGTGGGTGAA
ATCCTTGGCGTGACCTTCTGCGCGGCGTCGCAGGCTTTCTGGTGGGTTAA
-------------CCTTCT-------------------------------
ATCGTAGCTGTTACTTTCCCTGCGGCTGTGCAAGCTTTCGCATGGGGTCG
ATCGTTTCTATTACTTTCCCTGCTGCTGTTCAAGCTTGCTTGTGGTGGCG
ATTGTTGGTGTTACATTCCCTGCTGCCGTACAAGCCGTACTTTGGTATCG
--------------------------------------------------
ATCGTTGGTATTACTTTCCCAGCCGCTGTTCAAGCAGTTGTGTGGTATCG
ATCGTTGGTATTACTTTCCCAGCCGCTGTTCAAGCAGTTGTGTGGTATCG
ATCGTTGGTATTACATTCCCGGCTGCGGTTCAAGCAGTTCTTTGGTACCG
ATCGTTGGTATTACATTCCCAGCCGCTGTTCAAGCAGTTCTTTGGTACCG
ATCGTTGGTATTACATTCCCAGCCGCTGTTCAAGCAGTTCTTTGGTACGG
```

```
ATCGTGCTGGTAACTTTCCCGGCTGCGGTACAGGCCTTCTTGTGGGAGCG
ATCGTGTTGGTCACCTTCCCGGCCGCCACCCAGGTGATGCTGTGGGAGCG
ATCGTGTTGGTCACCTTCCCGGCTGCCGTTCAATACTTCCTGTGGGAACG
ATCGTGTTGGTGACGTTCCCGGCGGCGATCCAGGCGTGGACCTGGGAACG
ATCGTACTGGTCACCTTCCCGGCGGCCGTACAATCCTACCTGTGGGAGCG
ATCGTGTTGATCACGTTTCCGGCGGCGGTACAGGCGTACTTGTGGGAGCG
ATCGTGTTGATCACGTTCCCGGCGGCGGTACAGTCGTACCTGTGGGAGCG
ATTGTGATGATCACCTTCCCGGCGGCGGTGCAAGCGGTGCTGTGGGAGCG
ATCGTCATGATCACCTTCCCGGCCGCGGTGCAAGCGGTGATTTGGGAACG

CTTCCGCCTTCCTTTCGGCGCGACCTTCCTGGTCCTCGGCCTTCTCTTCG
TCTCCGCGTTCTGTTCGGCGCCGTTTTCGCGGCTCTGGGCCTCCTGATCG
TCTCCGCGTTCCGTTCGGCGCCGTTTTCGCGGCTCTGGGCCTTCTGATTG
CTTCCGTCTGCCGTTTGGCGCGGTGTTCGCGGCTCTTGGCCTTCTGATTG
CTTCCGTCTGCCGTTCGGCGCGGTGTTTGCGGCTCTCGGCCTTCTGATTG
CTTCCGTCTTCCGTTCGGCGCCGTTTTCGCGGCTCTGGGCCTGCTGATTG
CTTCCGCCTTCCGTTCGGCGCCGTGTTCGCGGCTCTGGGCCTCCTGATTG
CTTCCGTCTGCCGTTCGGCGCGGTGTTCGCGGCTCTCGGCCTGCTGATCG
CTTCCGCCTTCCGTTCGGCGCCGTTTTCGCGGCTCTTGGCCTCCTGATCG
TCTCCGCCTTCCGTTCGGCGCCGTTTTCGCGGCTCTGGGCCTCCTGATTG
TTACCGTTTGCCTTGGGGCGCAACCGTTTGCGTTCTGGGCTTGCTGTTCG
TTACCGTTTGCCATTCGGCGCTGTTGTTTGTATCTTAGGCTTGCTTTTGG
CTACCGTTTGCCGTTCGGTGCAGTTATTTCCGTTCTGGGTCTGCTGTTTG
----------------------------------------------TGG
TTGGCGTCTGCCATTCGGCGCAATGCTGGCTGTTCTGGGCCTGCTGTTCG
TTGGCGTCTGCCATTCGGCGCAATGCTGGCTGTTCTGGGCCTGCTGTTCG
TTACCGTCTGCCATTCGGCGCTATGCTGGCTGTATTAGGTCTGCTGTTCG
CTATCGTCTGCCATTCGGCGCAATGCTGGCTGTTCTTGGCCTGTTGTTCG
CTATCGTCTGCCATTCGGCGCAATGCTGGCTGTTCTTGGCCTGTTGTTCG
TTTCCGTCTGCCTTGGGGCGCGACGATTTGCGTACTGGGCCTGCTGTTTG
TCTGCGTCTTCCTTGGGGCGCGACGGTATGCGTATTGGCCCTGTTGTTTG
GTTCCGTCAGCCCTGGGGTGCGACCGTGTGCGTACTGGCGCTGCTCTTCG
GTTCCGCAACCCGTGGGGCGCAACCATGGCCGTTCTGGCGCTGCTCTTCG
GTATCGTCTGCCCTGGGGAGCCACCGTGTGCGTCCTGGGTCTGCTGCTGG
GTATCGTCAACCGTGGGGTGCGACGGTGTGCGTGCTGGGTCTGTTGTTGG
GTTTCGTCAACCGTGGGGTGCTACGGTCTGCGTACTGGGCCTGTTGATGG
TTTGCGGCTGCCGTGGGGGGCGACGGTCTGTGTTTTGGGCATCCTGTTTG
TCTGCGCCTGCCTTTCGGGGCCACCATTTCCATTCTGGGCATTCTGCTGG

GTGAATGGGTCAACCGCTACACCAATTTCTGGGGTTGGACCTATTTCCCG
GCGAGTGGATCAATCGCTACGTCAACTTCTGGGGCTGGACCTATTTCCCG
GCGAGTGGATTAACCGCTACGTCAACTTCTGGGGCTGGACGTATTTCCCG
GCGAGTGGATCAACCGCTACGTCAACTTCTGGGGCTGGACCTACTTCCCG
GCGAGTGGATCAACCGCTACGTCAACTTCTGGGGCTGGACCTACTTCCCG
GCGAGTGGATCAACCGCTACGTCAACTTCTGGGGCTGGACGTATTTCCCG
GCGAGTGGATCAACCGCTACGTCAACTTCTGGGGCTGGACGTACTTCCCG
GCGAGTGGATCAACCGCTACGTCAACTTCTGGGGCTGGACCTACTTCCCG
GCGAGTGGATCAACCGCTACGTCAACTTCTGGGGCTGGACCTACTTCCCG
GCGAGTGGATCAACCGCTACGTCAACTTCTGGGGCTGGACCTATTTCCCG
GTGAATGGGTTAACAGATACTTCAACTTCTGGGGTTGGACATACTTCCCG
GTGAGTGGATCAACAGATACTTGAACTTCTGGGGCTGGACTTACTTCCCT
GTGAGTGGGTTAACAGATACTTCAATTTCTGGGGATGGACTTATTTCCCA
GGGA--------------------ACTTCTGGGGGTGGACTTATTTCCCA
GCGAATGGGTTAACAGATATTTCAACTTCTGGGGATGGACTTACTTCCCA
GCGAATGGGTTAACAGATATTTCAACTTCTGGGGATGGACTTACTTCCCA
GCGAATGGGTCAACAGATACTTCAACTTCTGGGGATGGACTTACTTCCCG
GTGAATGGGTTAACAGATATTTCAACTTCTGGGGATGGACTTACTTCCCA
```

GTGAATGGGTTAACAGATATTTCAACTTCTGGGGATGGACTTACTTCCCA
GTGAGTGGGTCAACCGCTACTTCAACTTCTGGGGCTGGACCTACTTCCCG
GTGAGTGGATCAACCGCTACTTCAACTTCTGGGGTTGGACCTACTTCCCG
GTGAATGGGTCAACCGCTACTTTAACTTCTGGGGCTGGACCTACTTCCCC
GCGAATGGGTCAACCGCTACTTCAACTTCTGGGGCTGGACCTACTTCCCG
GCGAGTGGATCAACCGTTATTTCAACTTCTGGGGCTGGACCTACTTCCCG
GTGAGTGGGTCAACCGCTACTTCAACTTCTGGGGCTGGACGTATTTTCCG
GTGAGTGGATCAACCGTTACTTCAACTTCTGGGGCTGGACCTACTTCCCG
GGGAATGGGTGAACCGTTACTTTAACTTCTGGATGTGGACCTATTTCCCC
GTGAATGGATCAACCGCTACTTCAACTTCTGGGGATGGACCTACTTCCCG

ATCAGCCTCGTTTGGCCGACCTCTTTGGTCCCGGCTGCTCTGTTCCTCGA
ATCAGCCTCGTGTTCCCGTCCGCTCTGATCGTCCCGGCGATCTGGCTCGA
ATCAGCCTCGTTTTCCCGTCTGCCCTGATCGTTCCGGCGATCTGGCTTGA
ATCAGCCTTGTGTTCCCGTCTGCGTTGATCGTTCCGGCGATCTGGCTTGA
ATCAGCCTCGTGTTCCCGTCTGCTCTGATCGTTCCGGCGATCTGGCTCGA
ATCAGCCTCGTGTCCCCGTCCGCTCTGATCGTTCCGGCGATCTGGCTCGA
ATCAGCCTCGTGTTCCCGTCCGCTCTGATCGTTCCGGCGATCTGGCTCGA
ATCAGCCTGGTGTTCCCGTCTGCTCTGATCGTTCCGGCGATCTGGCTGGA
ATTTCGCTGGTGTTCCCGTCTGCTCTGATCGTTCCGGCGATCTGGCTCGA
ATCAGCCTCGTGTTCCCGTCCGCTCTGATCGTTCCGGCGATCTGGCTTGA
GTTAACTTCGTATTCCCATCACAATTGATTCCTAGCGCTATCCTGCTGGA
GTTAACTTCGTGTTCCCATCAAACTTGATGCCAGGCGCTATCATTCTTGA
GTGAATTTCGTATTCCCATCACAATTCGTGCCAGGCGCAATCGTTCTTGG
GTAAACTTCGTGTTCCCTTCAAACTTCATGCCTGGCGCTATCGTTCTGGA
GTTAACTTCGTGTTCCCATCACAATTCGCTCCAGGCGCAATCGTTCTGGA
GTTAACTTCGTGTTCCCATCACAATTCGTTCCAGGCGCAATCGTTCTGGA
GTTAACTTTGTATTCCCATCACAATTCGTTCCAGGCGCACTCGTTCTGGA
GTTAACTTTGTATTCCCATCACAATTCGTTCCAGGCGCACTGGTTCTGGA
GTTAACTTTGTATTCCCATCACAATTCGTTCCAGGCGCACTGGTTCTGGA
ATCACCCTGTGCTTCCCGTCGCAGATCGTACCGGGCGCCATCCTGCTCGA
ATCACCCTGTGCTTCCCCTCGCAGATCGTGCCGGGCGCCATCCTGCTCGA
GTGAACTTCGTGTTCCCGGCGGTCCTGGTTCCGGGCGCGATCCTGCTCGA
ATCAACTTCGTATTCCCGGCCATTCTGGTTCCGGGTGCGATCTTGCTCGA
ATCAACTTCGTGTTCCCTGCCTCGCTGGTGCCGGGCGCCATCATCCTGGA
GTGGACTTCGTATTTCCGGCGATTTTGACGCCGGGTGCGATTTTGCTGGA
GTCAATTTCGTGTTCCCGGCGATTCTGACCCCGGGCGCGATCCTGCTCGA
ATTAACTTTGTTTTCCCGACGGCGGCGGTGCACATGGCGATCTTTTTAGA
ATCAACTTCGTCTTCCCGACGGCGGCGGTCCACATGGCCATCTTCCTGGA

TATCGTGCTCTTGCTGAGCCGCAGCTTCATCGTCACCGCGATCGTCGGCG
CGTGATCCTGCTTCTGTCGGGTTCCTATGTGATCACGGCGGTTGTCGGTT
CGTGATCCTGCTCCTGTCGGGTTCCTATGTGATCACGGCGGTTGTCGGTT
CGTGATCCTGCTTCTGTCGGGGTCCTATGTGATCACGGCGATTGTTGGTT
CGTGATCCTGCTTCTGTCGGGCTCCTATGTGATCACGGCGGTTGTCGGCT
CGTGATCCTGCTGCTGTCGGGCTCCTATGTGATCACGGCGGTTGTCGGTT
CGTGATCCTGCTTCTGTCGGGTTCCTATGTGATCACGGCGGTTGTCGGTT
CGTGATCCTGCTGCTGTCGGGCTCCTATGTGATCACGGCGGTTGTCGGTT
CGTGATCCTGCTGCTGTCGGGCTCCTATGTGATCACGGCGGTTGTCGGTT
CGTGATCCTGCTTCTGTCGGGTTCCTATGTGATCACGGCGGTTGTCGGTT
CGTGGTTCTGTTGCTGTCCAACAGCTACACCTTCACCGCTGTTGCTGGCG
TGTTGTATTGTTGTTGTCAAACAGCATGGTATTGACTGCCGTTATCGGTG
TGTTATCCTGATGCTGTCCAACAGCATGCAATTAACAGCCGTTATCGGTG
CGTTATCCTGATGCTGTCCAACAGCATGCAGCTGACGGCTGTTATCGGTG
CGTAATTCTGATGCTGTCTAACAGCATGCAGTTGACTGCGGTTCTGGGCG
CGTAATTCTGATGCTGTCTAACAGCATGCAGTTGACTGCGGTTCTGGGCG
CGTGATCCTGATGTTGTCAAACAGCATGCAGTTGACTGCTGTGCTGGGCG

CGTAATTCTGATGCTGTCTAACAGCATGCAGTTGACTGCTGTTATCGGCG
CGTAATTCTGATGCTGTCTAACAGCATGCAGTTGACTGCTGTTATCGGCG
CACGGCTCTGATGCTGAGCACCAGCTACCTGTTCACCGCGATCGTCGGTG
CGTCGTTCTGTTGCTCTCCGGCAGCTACCTGCTCACCGCCATCGTGGGCG
CGTTATCCTGATGCTCTCGGGCAGCTACCTGTTTGCTGCCATCATCGGCG
CACCTTCCTGATGCTTTCGGGCAGCTACCTGTTCACGGCGATCGTGGGCG
CACCGTGCTGATGCTGTCGGGCAGCTACCTGTTCACCGCGATCGTCGGTG
CACGATGCTGTTGCTGAGCGGCAGCTATCTGTTCACGGCGATCATTGGCG
CACGATGTTGATGCTGAGCGGCAGCTACCTGTTCACGGCGATCGTTGGCG
CGTGGTGCTGATGCTGTCTGGGAGTTTCTTGTTCACGGCGGTGATTGGCG
CGTGGTCCTGATGCTGTCCAGCAGCTTCCTGTTCACGGCCGTGGTTGGCG

CGATGGGCTGGGGTCTTCTCCTCTATCCCTCCAACTGGCCGATTCTTGCG
CGCTGGGCTGGGGTCTGCTGTTCTACCCGAACAACTGGCCGGCGATTGCG
CGCTGGGCTGGGGTCTGCTGTTCTACCCGAACAACTGGCCGGCGATCGCC
CGCTGGGCTGGGGTCTGTTGTTCTACCCGAACAACTGGCCGGCGATTGCG
CGCTGGGCTGGGGTCTGCTGTTCTATCCGAACAACTGGCCGGCGATCGCC
CGCTGGGCTGGGGTCTGCTGTTCTATCCGAACAGCTGGCCGGCGATTGCG
CGCTGGGCTGGGGTCTGCTGTTCTACCCGAACAACTGGCCGGCGATTGCG
CGCTGGGCTGGGGTCTGCTGTTCTATCCGAACAACTGGCCGGCGATCGCC
CGCTGGGCTGGGGTCTGCTGTTCTACCCGAACAACTGGCCGGCGATCGCC
CGCTGGGCTGGGGTCTGCTGTTCTACCCGAACAACTGGCCGGCGATCGCC
CTATGGGTTGGGGCTTGATTTTCTATCCTAGCAACTGGCCGGTTATTGGT
GTTTGGCTTATGGTTTGGTTTTCTACCCAGGCAACTGGCCAATCATTGCT
GTTTGGGCTATGGTCTGTTGTTCTATCCTGGCAACTGGCCTGTCATCGCT
GTTTAGGCTACGGCTTGTTGTTCTATCCTGGTAACTGGCCAGTCATCGCT
GCTTGGCTTATGGGCTGTTGTTCTATCCTGGCAACTGGCCGGTCATCGCT
GCTTGGCTTATGGCCTGTTGTTCTATCCTGGCAACTGGCCGGTCATCGCT
GCTTGGCTTATGGTCTGTTGTTCTATCCTGGCAACTGGCCTGTCATCGCT
GTCTGGCATACGGCTTGTTGTTCTATCCTGGCAACTGGCCTGTCATCGCT
GTCTGGCATACGGCTTGTTGTTCTATCCTGGCAACTGGCCTGTCATCGCT
CGATGGCTTGGGGCTTGGTTTTCTATCCGGGCAACTGGCCGGTGATTGCA
GCATGGCTTGGGGCTTGATCTTCTACCCGGGCAACTGGCCGGTGATCGCG
GTCTGGCTTGGGGCCTGATTTTCTATCCGGGCAACTGGCCGGTGATTGCT
GCATGGCTTGGGGCCTGATTTTCTATCCGGGCAACTGGCCGATGATCGCA
CGATGGGCTGGGGTCTGACCTTCTACCCGGGCAACTGGCCGATCATCGCG
GTTTGGGCTGGGGTTTGATTTTCTACCCGGGCAACTGGCCGATCATTGCG
CGATGGGTTGGGGCCTGATTTTCTATCCGGGCAACTGGCCGATCATTGCA
GGTTAGGCTGGGGTTTGTTGATGTATCCTGGCAACTGGCCGATTATTGCG
GCTTGGGCTGGGGTCTGTTGATGTACCCGGGCAACTGGCCGGTGATCGCA

CCTTATCATCAGGCGACGGAACAATATGGCCTTCTCATGTCTCTCGCCGA
GCGTTCCACCAGGCGACCGAGCAGCATGGTCAGCTGATGACGCTTGCTGA
GCCTTCCACCAGGCGACCGAGCAGCATGGTCAGCTGATGACCCTCGCCGA
GCGTTCCACCAGGCGACGGAGCAGCATGGTCAGCTGATGACGCTTGCGGA
GCCTTCCACCAGGCGACCGAGCAGCATGGTCAGCTGATGACGCTGGCCGA
GCGTTCCACCAGGCGACTGAGCAGCACGGCCAGCTGATGACGCTTGCTGA
GCGTTCCACCAGGCGACCGAGCAGCATGGTCAGCTGATGACCCTGGCTGA
GCCTTCCACCAGGCGACCGAGCAGCATGGTCAGCTGATGACGCTTGCTGA
GCCTTCCACCAGGCGACCGAGCAGCATGGTCAGCTGATGACCCTGGCTGA
GCCTTCCACCAGGCGACCGAGCAGCATGGTCAGCTGATGACCCTGGCTGA
CCATTACACGTGCCTGTTGAATATAACGGCATGATGATGACTTTGGCTGA
CCATTGCACGTTCCTGTTGAATACAATGGCATGATGATGACAATTGCTGA
CCATTGCACGTGCCTGTTGAATATAACGGTATGGTTATGACGCTGGCTGA
CCATTGCACGTCCCTGTTGAATACAATGGTATGGTAATGACTCTGGCTGA
CCATTGCACGTGCCAGTTGAATACAACGGGATGGTAATGACCCTGGCTGA
CCATTGCACGTGCCAGTTGAATACAACGGCATGGTAATGACCCTGGCTGA

```
CCATTGCACGTGCCAGTTGAATACGACGGCATGGTAATGACCCTGGCTGA
CCATTGCACGTGCCTGTTGAATACAACGGCATGGTAATGACGCTGGCTGA
CCATTGCACGTGCCTGTTGAATACAACGGCATGGTAATGACGCTGGCTGA
CCGTACCACGTGCCGGTAGAGTACAACGGCATGCTGATGTCGGTTGCCGA
CCGTACCACGTGCCGGTGGAATACAACGGCATGCTGCTGTCGGTGGCCGA
CCGCTGCACGTGCCGGTGGAATACAACGGCATGTTGATGTCCATTGCCGA
CCGCTGCACGTGCCGGTGGAATACAACGGCATGCTCATGTCGATTGCCGA
CCGCTGCACGTGCCGGTGGAATACAACGGCATGCTGATGTCGATCGCCGA
CCGCTGCACGTACCGGTGGAGTACAGCGGGATGCTGATGTCGATTGCCGA
CCGATTCACGTGCCGGTGGAATACAGCGGCATGCTGATGTCGATTGCCGA
CCTTTGCACGTGCCGGTGGAATACAACGGCATGCTGATGTCGATTGCCGA
CCGCTGCACGTGCCTGTGGAATACAACGGCATGCTGATGTCCGTGGCCGA

CT-TGATTGGCTTTGAATACGTCCGCACC-TCGATGCCT-GAATATCTCC
CC-TGATCGGCCTCCACTACGTCCGCACG-TCGATGCCG-GAATACATCC
CC-TGATTGGTCTGCACTACGTCCGCACG-TCGATGCCG-GAATACATCC
TC-TGATCGGCTTCCACTTCGTGCGCACGCTCGATGCCG-GAATATATCC
CC-TCATCGGCTTCCACTTCGTCCGCACC-TCGATGCCG-GAATACATCC
CC-TGATCGGCCTCCACTACGTCCGCACG-TCGATGCCG-GAATACATCC
CC-TGATCGGCCTTCACTGTGTCCGCACG-TCGAAGCCG-GAATACATCC
CC-TGATCGGCCTCCACTACGTCCGCACG-TCGATGCCG-GAATACATCC
CTCTGATCGGCCTCCACTACGTCCGCACG-TCGATGCCG-GAATACATCC
CC-TGATCGGCCTCCACTACGTCCGCACG-TCGATGCCG-GAATACATCC
CT-TACAAGGTTACCACTATGTAAGAACC-GGTACTCCT-GAATACATCC
CT-TACAAGGTTACCACTACGTAAGAACT-GGTACACCT-GAGTACATCA
CT-TGCAAGGTTACCACTATGTAAGAACC-GGTACCCCA-GAGTACATCA
CT-TGCAAGGTTACCACTATGTAAGAACT-GGTACACCT-GAATACATCC
CT-TGCAAGGTTACCACTATGTAAGAACC-GGTACTCCA-GAATATATCC
CT-TGCAAGGTTACCACTATGTAAGAACC-GGTACTCCA-GAATATATCC
CT-TGCAAGGTTACCACTACGTAAGAACT-GGTACTCCA-GAGTACATCA
CT-TGCAAGGTTACCACTATGTAAGAACT-GGTACTCCA-GAATACATCA
CT-TGCAAGGTTACCACTATGTAAGAACT-GGTACTCCA-GAATACATCA
CC-TGCTGGGTTACCACTATGTCCGTACG-GGTACGCCT-GAGTACATCC
CT-TGCTGGGCTACAACTATGTTCGTACC-GGCACCCCC-GAGTACATCC
CA-TCCAGGGCTACAACTATGTTCGTACC-GGTACCCCC-GAGTACATCC
CT-TGCAGGGCTACCACTATGTCCGTACC-GGTACCCCGCGAGTACATCC
CA-TCCAGGGTTACAACTATGTGCGTACG-GGTACGCCT-GAGTACATCC
CA-TTCAGGGTTACAACTACGTGCGTACC-GGAACCCCG-GAATACATCC
CA-TTCAGGGTTACAACTACGTGCGTACCTGGTACGCCG-GAATACATCC
CA-TTCAAGGCTACCACTACGTGCGCACG-GGGACGCCG-GAATACATTC
CA-TCCAAGGCTACCATTACGTCCGTACG-GGTACGCCC-GAATACATCC

GCATCGTCGAGCGCGGCACGATGCGCACCTTCGGTAAGGACGTCGTCGCG
GCATGGTCGAGCGCGGCACGCTGCGCACGTTCGGTAAGGACGTTGCTCCG
GCATGGTCGAGCGCGGCACGCTGCGCACGTTCGGTAAAGACGTTGCTCCG
GCAT-GTCGAGC--------------------------------------
GCATGGTCGAGCGCGGCACGCTGCGCACCTTCGGTAAAGACGTTGCGCCG
GCATGGTCGAGCGCGGCACGCTGCGCACGTTCGGTAAAGACGTTGCTCCG
GCATGGTCGAGCGCGGCACGCTGCGCACGTTCGGTAAAGACGTTGCTCCG
GCATGGTCGAGCGCGGCACGCTGCGCACGTTCGGTAAAGACGTTGCGCCG
GCAT-GTCGAGCGCGTA----------------------------------
GCATGGTCGAGCGCGGCACGCTGCGCACGTTCGGTAAGGACGTTGCTCCG
GTATGGTTGAAAAAGGTACATTGAGAACTTTCGGTAAAGACGTTGCTCCG
GAATGGTTGAAAAAGGTACCTTAAGAACATTCGGTAAAGACGTTGCTCCA
GAATGGTTGAAAAAGGTACGTTGAGAACATTTGGTAAAGACGTTGCGCCG
GGATGGTTGAAAAAGGTACACTGAGAACTTTCGGTAACGACGTCGCCCCC
GGATGGTTGAAAAAGGTACTCTGAGAACTTTCGGTAAGGACGTTGCTCCA
```

```
GGATGGTTGAAAAAGGTACTCTGAGAACTTTCGGTAAGGACGTTGCTCCA
GAATGGTTGAAAAAGGTACTTTGAGAACATTCGGTAAAGACGTTGCTCCG
GAATGGTTGAAAAAGGTACTTTGAGAACTTTCGGTAAGGACGTTGCTCCG
GAATGGTTGAAAAAGGTACTTTGAGAACTTTCGGTAAGGACGTTGCGCCG
GTATGGTCGAGAAGGGCACCCTGCGTACCTTCGGTAAAGACGTTGCTCCG
GCATGGTTGGGAAAGGCACCCTGCGTACCTTCGGTAAAGACGTTGCGCCG
GCATGGTTGAGAAAGGCACCCTGCGTACCTTCGGTAAAGACGTTGCTCCG
CGC-----------------------------------------------
GCATGGTAGAGAAGGGCACCCTGCGTACCTTCGGTAAGGAC---------
GGATGGTCGAGAAGGGCACGCTGCGTACGTTTGGTAAAGACGTTGCGCCG
GCA--GCTGAGAA-------------------------------------
GGATGGTAGAGAAAGGCACGCTGCGTACCTTTGGTAAGGACA--------
GGATGGTTGAAAAAGGCACCCTGCGTACCTTCGGTAAAGACGTGGCGCCG

G-----------------------------TTGCGGCCTTCT-------
GAAGGGCGAAT----------TCCAGCACACTGGCGGCCGTTA-------
GAAGGGCGAAT----------TCCAGCACACTGGCGGCCGTTA-------
-----------------------------------------------
GAAGGGCGAATTCTGCAGATATCCATCACACTGGCGGCCGCTCGAGCATG
GAAGGGCGAATTCTGCAGATATCCATCACACTGGCGGCCGCTCGAGCATG
GAAGGGCGAATTCTGCAGATATCCATCACACTGGCGGCCGCTCGAGCATG
GAAGGGCGAATTCTGCAGATATCCATCACACTGGCGGCCGCTCGAGCATG
-----------------------------------------------
GAAGGGCGAATTCTGCAGATATCCATCACACTGGCGGCCGCTCGAGCATG
GAAGGGCGAAT----------TCCAGCACACTGGCGGCCGTTA-------
GTA------------------TCAGCCTTCTTCTCAGCCTTC--------
GAAGGGCGAATTCTGCAGATATCCATCACACTGGCGGCCGCTCGAGCATG
AG-----------------------------------------------
GTA------------------TCC--------------------------
GTA------------------TCTGCCTTCTTCTCCGCCTTC--------
GAAGGGCGAAT----------TCCAGCACACTGGCGGCCGTTA-------
GAAGGGCGAAT----------TCCAGCACACTGGCGGCCGTTA-------
GAAGGGCGAAT----------TCCAGCACACTGGCGGCCGTTA-------
GAAGGGCGAATTCTGCAGATATCCATCACACTGGCGGCCGCTCGAGCATG
GAAGGGCGAAT----------TCCAGCACACTGGCGGCCGTTA-------
GAAGGGCGAAT----------TCCAGCACACTGGCGGCCGTTA-------
-----------------------------------------------
-----------------------------------------------
GAAGGGCGAAT----------TCCAGCACACTGGCGGCCGTTA-------
-----------------------------------------------
-----------------------------------------------
GTA------------------TCC--------------------------

-----------TCTC------AGCC-------------------------
---CTAGTGGATCCG------AGCTCGGTACCAAGCTTGGCGTAATCATG
---CTAGTGGATCCG------AGCTCGGTACCAAGCTTGGCGTAATCATG
-----------------------------------------------
CATCTAGAGGGCCCANNNANNNNNCTCGTNCTANNNNN------------
CATCTAGAGGGCCNG------AGNGNAACAGCTAGANCGANNNN------
CATCTAGAGGGCCCCNATNAAACAGNATAGNNGANNN-------------
CATCTAGAGGGCCCNNNNNACAGCTAGTAGANNNN---------------
-----------------------------------------------
CATCTAGAGGGCCC---------CNTGNANANCAGTATAGGACNATNNNN
---CTAGTGGATCCG------AGCTCGGTACCAAGCTTGGCGTAATCATG
-----------------------------------------------
CATCTAGAGGGCCCA-----------NNANGNCAGTATAGNACANNNN--
-----------------------------------------------
```

```
--------------------------------------------------
--------------------------------------------------
---CTAGTGGATCCG------AGCTCGGTACCAAGCTTGGCGTAATCATG
---CTAGTGGATCCG------AGCTCGGTACCAAGCTTGGCGTAATCATG
---CTAGTGGATCCG------AGCTCGGTACCAAGCTTGGCGTAATCATG
CATCTAGAGGGCCCN------NNNNNANCAGCTAGTANNTANNN------
---CTAGTGGATCCG------AGCTCGGTACCAAGCTTGGCGTAATCATG
---CTAGTGGATCCG------AGCTCGGTACCAAGCTTGGCGTAATCATG
--------------------------------------------------
--------------------------------------------------
---CTAGTGGATCCG------AGCTCGGTACCAAGCTTGGCGTAATCATG
--------------------------------------------------
--------------------------------------------------
--------------------------------------------------

--------------------------------------------------
GTCATGNTTGTTTTCCAAAANNCGGTNCGAACCCTGCTCGCTACCTAGCT
GTCATAGCTGATTCCAA--------------------------------
--------------------------------------------------
--------------------------------------------------
--------------------------------------------------
--------------------------------------------------
--------------------------------------------------
--------------------------------------------------
--------------------------------------------------
GTCATAGCTTGTTTCCANN------------------------------
--------------------------------------------------
--------------------------------------------------
--------------------------------------------------
--------------------------------------------------
--------------------------------------------------
GTCATNCCATGTTTTCCNA------------------------------
NNATNCTTATTTTCAA---------------------------------
TNAANNTTTTTTTNCCNN-------------------------------
--------------------------------------------------
GTCAANNTTTTTTTCAA--------------------------------
GTCAANNCTGTTTTCNA--------------------------------
--------------------------------------------------
--------------------------------------------------
GTCATAGCCTGTTTCN---------------------------------
--------------------------------------------------
--------------------------------------------------
--------------------------------------------------

--------------------------------------------------
CGCCGCAATCGTGTTCTGGTATGTTTCCAAAATTANNAATTATGTTTTTA
--------------------------------------------------
--------------------------------------------------
--------------------------------------------------
--------------------------------------------------
--------------------------------------------------
--------------------------------------------------
```

Appendix 14

Phylogenetic tree in Newick format

((((((((((Uncultured:0.0353195621,12b:0.1121838767)96:0.0613102849,Methylomon:0.1008367951)72:0.0415806968,(Methylocal:0.1300937628,13a:0.0548511424)53:0.0238876530)87:0.0360449527,(Methylohal:0.0833197835,Methylothe:0.1671028379)100:0.1633741749)58:0.0197592685,(((10b:0.0166862811,(13d:0.0345141444,(4:0.0081192040,((Methylosin:0.0262043540,((Methylocys:0.0536273270,16b:0.0275261644)61:0.0135176853,16c:0.0116954369)56:0.0242964028)32:0.0076324290,(15i:0.0413075355,15e:0.0212838344)42:0.0174372536)51:0.0129518717)76:0.0433967418)31:0.0092357072)100:0.2113998169,Methanocap:0.3611779307)100:0.2102400020,(15b:0.0918625762,10c:0.1107440010)84:0.0285372714)79:0.0217599827)100:0.2087432976,16e:0.0388172279)100:0.0988890511,Methylosom:0.2040444839)53:0.0434871805,Methylosar:0.1237491794)37:0.0192990011,12c:0.1373471236)87:0.0482834413,(15d:0.0075289303,13b:0.0055459195)97:0.0280991851)41:0.0132175316,10a:0.0465773782)100:0.0731975070,Methylomic:0.0029222388,Methylobac:0.0064395860);