PCR PRIMERS FOR THE DETECTION OF PROPANE AND BUTANE-OXIDIZING MICROORGANISMS

A Thesis

presented to

the Faculty of California Polytechnic State University,

San Luis Obispo

In Partial Fulfillment

of the Requirements for the Degree

Master of Science in Biological Sciences

by

Brian Jeremy Chan

March 2011

COMMITTEE MEMBERSHIP

TITLE: PCR PRIMERS FOR THE DETECTION OF PROPANE AND BUTANE-OXIDIZING MICROORGANISMS

AUTHOR: Brian Jeremy Chan

DATE SUBMITTED: March 2011

Committee Chair: Christopher Kitts, Ph.D.

Committee Member: Raul Cano, Ph.D.

Committee Member: Pat Fidopiastis, Ph.D.

ABSTRACT

PCR Primers for The Detection of Propane and Butane-Oxidizing Microorganisms

Brian Jeremy Chan

In an increasingly energy-hungry world, our capacity to meet the heightened energy demands of the future has become a pressing matter. The most urgent of these concerns are tied to the accessibility of petroleum. Various experts have proselytized both the imminent arrival of peak oil production rates and the ensuing decline of those rates thereafter. And to that end, the development of novel and advanced oil exploration methodologies has become almost as important as finding the sources of oil themselves.

The soils above petroleum reservoirs play host to various communities of alkane-oxidizing bacteria that can utilize the natural gas emitted by the reservoirs as a source of carbon and energy. While methane can originate from non-petroleum sources, the only natural sources of propane and butane are oil and gas fields. The increased presence of propane and butane-oxidizing bacteria in a given soil sample is used by oil prospectors as an accurate indicator of a proximal petroleum reservoirs.

For over a century, cell counts and hydrocarbon metabolic rates have been the metrics used to determine the presence of hydrocarbon-oxidizing microbes. These methods require weeks to complete. Here, we have developed a set of DNA primers for a much more rapid detection of hydrocarbon-oxidizing microbes through PCR amplification - for the chief purpose of petroleum exploration. Each primer's design is based on a nucleotide sequence alignment of seven *prm*A and *bmo*X genes from seven organisms, which encode the large hydroxylase subunit of propane monooxygenase and

the alpha hydroxylase subunit of butane monooxygenase respectively. These monooxygenases are the enzymes responsible for the initiation of propane and butane catabolism. Optimization of PCR with this primer set was accomplished using DNA extracted from known butane and propane oxidizers as positive controls, and methane and toluene oxidizers as negative controls. PCR products recovered from cultures of butane-oxidizing and propane-oxidizing bacteria, and soil samples, were sequenced. Phylogenetic trees were constructed from the sequencing data to confirm the accuracy of amplification. We demonstrate the use of PCR and agarose gel electrophoresis to detect hydrocarbon-oxidizing bacteria in culture and in complex microbial soil communities. Detection limits were elucidated through two different experiments. Potential avenues of advancements include narrowing specificity by selectively removing primer degeneracies, the use of additional positive and negative controls and the adaptation of the primers to a qPCR TaqMan assay.

ACKNOWLEDGEMENTS

Dr. Christopher Kitts is responsible for much of my personal and professional development. He brought me into the Environmental Biotechnology Institute as an undergraduate, as an unknown quantity. Through the setbacks, his support never faltered. I am forever grateful for his encouragement at a time when all hope was lost. Dr. Raul Cano and Dr. Pat Fidopiastis' seemingly endless patience and insight were invaluable over the course of the thesis, from beginning to end. Our collaboration was a great learning experience. In addition to the committee, I'd like to thank Alice Hamrick for all the time and energy she afforded me over the years. She taught me basic lab skills and techniques that I will always carry with me. Finally, I'd like to thank my family and friends for supporting me over the years. Without them, I would not and could not have persevered.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1
INTRODUCTION

Petroleum Reservoirs

Oil is the lifeblood of modern civilization. It fuels the world's transportation and is the precursor to numerous petroleum-based products. But as energy consumption continues to skyrocket across the globe, supply will soon no longer be capable of satisfying world demand (Hirsch *et al*., 2006). By the 1980's, annual world oil consumption had already surpassed annual oil additions (Aleklett and Campbell, 2003). The rise of large industrializing nations has increased the demand for petroleum beyond previous forecasts (Streifel, 2006). This voracious appetite for petroleum has prompted the pursuit to engineer faster, more accurate, low-impact and cost-effective methodologies of prospecting. These advances while mostly geologically and chemically based, include biologically based methodologies as well.

Petroleum reservoirs are subsurface pools of short-chain and long-chain (C_1-C_{30}), linear, cyclic and branched alkanes (Hamamura *et al.*, 1999). Alkanes are a subset of hydrocarbons that possess only single covalent bonds. Petroleum, the liquid portion of alkanes, contains within it natural gases C_1-C_5 (McCain, 1990). Natural gas seeps from oil reservoirs to the earth's surface where it can enter into the earth's atmosphere. A fraction of the natural gas seepage is consumed by hydrocarbon-oxidizing (HCO) microorganisms that reside in the soil (van Beilen *et al.*, 2003). These HCO microorganisms can utilize short-chain alkanes (C_1-C_4) as their sole source of energy and carbon (Wagner *et al.*, 2002). They are of particular interest in petroleum prospecting because of their increased presence above oil reservoirs (Hitzman *et al.*, 2002).

Microbiology of Petroleum Reservoirs

Since the 1930's, alkane-oxidizing bacteria have been used to detect the presence of undiscovered petroleum reservoirs. Microbiologists G.A. Mogilewskii (1938) in the U.S.S.R. and M.S. Taggart (1941) and L.W. Blau (1942) in the United States all described the use of measuring concentrations of HCO microorganisms, in surface soil samples, as indicators of oil and gas fields in the deeper subsurface (Wagner *et al.*, 2002).

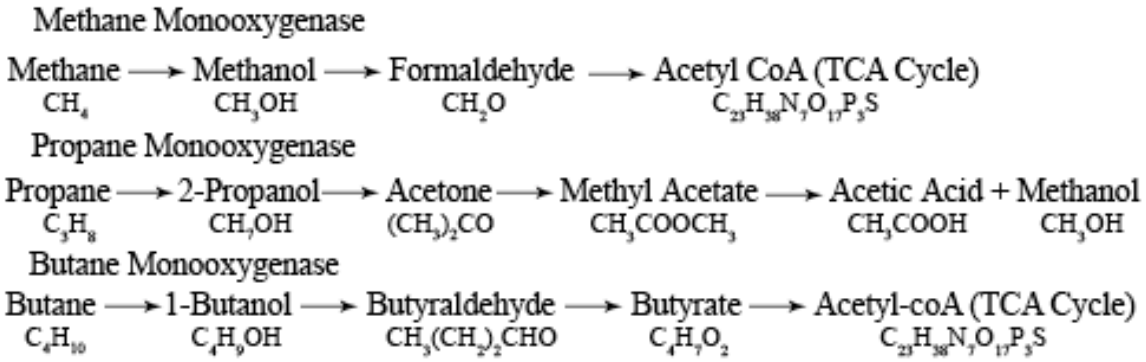

Figure 1. Metabolic pathways of methane, propane and butane oxidation, initiated by monooxygenases.

Natural gas contains methane (C_1, where the sub-index indicates the number of carbon atoms of the alkane molecule), ethane (C_2), propane (C_3), butane (C_4), and pentane (C_5) (Kinnaman *et al.*, 2007). In the soils above natural gas seepages, different groups of bacteria oxidize the various gases (Theisen and Murrell, 2005) (van Beilen *et al.*, 2003). With only a few exceptions, methane-oxidizing bacteria utilize methane exclusively as a carbon and energy source (Theisen and Murrell, 2005). The first step in the process, the oxidation of methane to methanol, is facilitated by a methane monooxygenase (MMO) (Fig. 1) (Colby and Dalton, 1976). After a series of reactions, methanol is eventually converted into either acetyl-coA and oxidized for energy or

converted into formaldehyde and assimilated by the cell (Higgins *et al.*, 1981). Conversely, propanotrophs and butanotrophs are facultative oxidizers of alkanes. In addition to hydrocarbons, they are able to utilize monosaccharides and polysaccharides for carbon and energy as well (Wagner *et al.*, 2002). Nevertheless, like methanotrophs, butanotrophs and propanotrophs initiate alkane metabolism with the oxidation of alkanes to alcohols via monooxygenase (Arp, 1999) (Kotani *et al.*, 2007).

Even though methane constitutes the greater part of natural gas, its presence does not always signify the presence of a petroleum reservoir or gas field. Melting methane clathrates and biogenic methanogenesis can also generate methane. Methane clathrates are lattices of frozen water that contain methane trapped within them. In response to anthropogenic global warming, an increasing number of subterranean methane clathrates are melting, releasing methane into the soil and atmosphere (Archer, 2007). Methanogenesis, the process of producing methane biogenically, involves a multispecies consortium of anaerobic bacteria and archaea that convert biomass into methane (Stolyar *et al.*, 2007). Molecular hydrogen and acetic acid released by heterotrophic anaerobes during the fermentation of organic compounds are used by methanogenic archaea as electron donors in the reduction of CO_2 to methane. Because of the multifactorial nature of methane formation, methanotrophs by themselves are not reliable indicators of petroleum reservoirs (Brisbane *et al.*, 1965).

Propane and butane are much more accurate indicators of petroleum reservoirs than methane. Large accumulations of butane and propane are limited to natural gas and petroleum reservoirs. In some cases, propane and butane can make up to 22% of the natural gas in an oil field (McCain, 1990). But as previously mentioned, propane and

butane do not represent exclusive energy or carbon sources for HCO microbes. Therefore, the mere presences of propanotrophs and butanotrophs cannot serve as indicators of petroleum. However, a constant supply of gaseous hydrocarbons can often lead to anomalously large populations of both communities. Therefore, anomalously large propane and butane-oxidizing populations can serve as accurate indicators of petroleum or natural gas in the subsurface. Anomalously large populations of C_2-C_4 oxidizers (10^4-10^6 cells • gram soil^{-1}) are several orders of magnitude larger than background C_2-C_4 oxidizer populations (10^3 cells • gram soil^{-1}) (Wagner *et al.*, 2002).

Microbial Oil Survey Technique

Microbial Oil Survey Technique (MOST) is a method of petroleum prospecting that is based on the detection of anomalously large HCO microbial populations above hydrocarbon seeps. Populations are characterized by cell counting, using a selective hydrocarbon medium. MOST samples are processed to identify the presence of C_2-C_4 oxidizing microbes, since there is a direct, positive relationship between the hydrocarbon concentrations in soil and these HCO microbial populations (Hitzman *et al.*, 2002).

The process begins with a soil dilution that is plated on a solid growth medium with n-butanol as its only carbon source. This selective growth medium allows only microorganisms with a light-hydrocarbon metabolism to survive. After a week of incubation, microorganisms grow into visible colonies and are counted. Counts are used to determine patterns of HCO densities that help define the hydrocarbon potential of an area. Geophysical and geologic data are used with MOST microbial seepage data to enhance petroleum exploration and production programs.

Microbial Prospecting for Oil and Gas

Microbial Prospecting for Oil and Gas (MPOG) is another microbiological technique that is currently used to explore for oil and gas. Soil samples are taken from a prospecting site, dispersed in a defined nutrient solution, and serially diluted. Separate replicates of each dilution are spiked with methane and a mixture light hydrocarbons and incubated at 30°C for 48 hours (Wagner *et al.*, 2002). Only those bacteria able to oxidize methane or light hydrocarbons in a short period will grow and consume a certain amount of supplied hydrocarbons. Cell concentration is estimated by Most Probable Number (MPN) and biochemical activity. Areal maps drawn with cell concentration data depict the distribution of methane and light hydrocarbon-oxidizing bacteria (Fig. 2). The higher the calculated and measured cell numbers and activity, the more intense the hydrocarbon seepage in the investigation area. Increased seepages over a site are interpreted as indicators of a subterrestrial hydrocarbon field in the vicinity.

Molecular Versus Culturing Techniques

Microbiology-based methodologies for petroleum prospecting employ different combinations of culturing, cell counting and metabolic activity, requiring weeks to produce results. They also require the use of incubators, media, and hydrocarbons. Conventional culturing methods are laborious and time consuming. In addition, more fastidious organisms may not grow in the artificial media provided, so counts could be an underestimation of actual cell numbers in the soil.

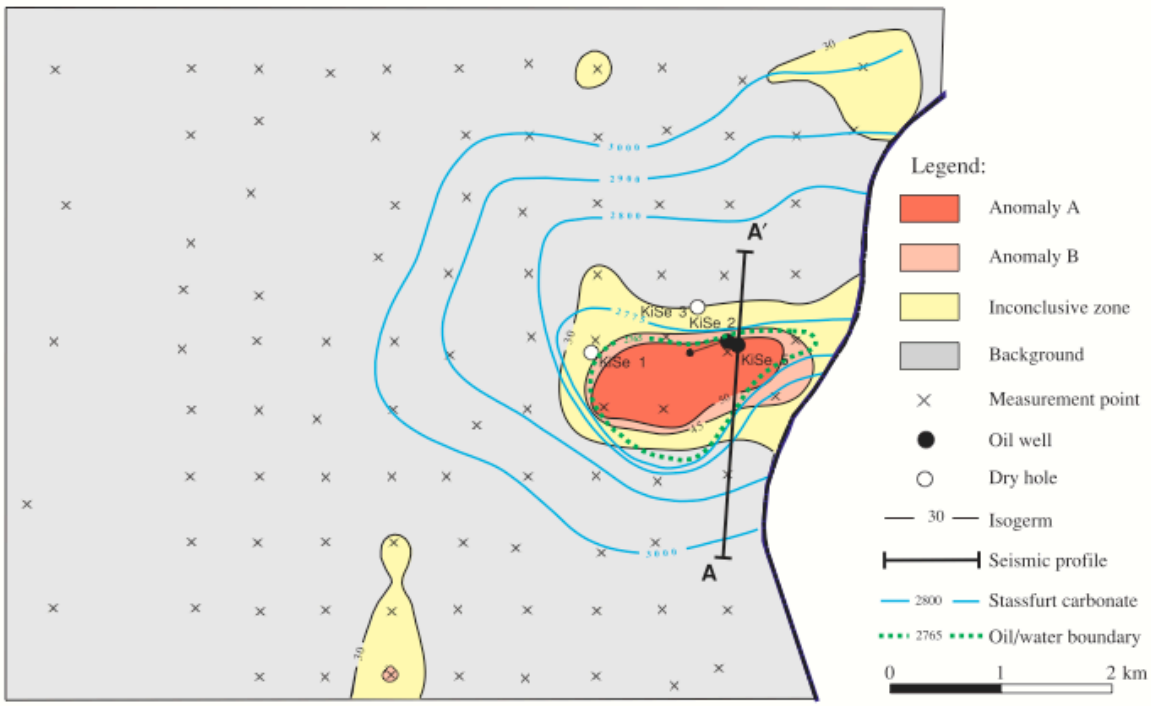

Figure 2. Example areal MPOG map of Kietz oil field, in Kietz, Germany.

The ability to detect a population of bacteria without the use of culturing is ideal. Molecular techniques provide a level of expedition and specificity that culturing cannot. PCR can provide a method of detecting bacteria and other organisms within a matter of hours. Molecular techniques aimed at detecting methanotrophic bacteria are currently available. One of these techniques utilizes PCR to amplify a portion of the *mmo*X gene (McDonald *et al*., 1995), which encodes the alpha subunit of hydroxylase of the soluble methane monooxygenase (sMMO). sMMO is an enzyme responsible for the oxidation of the C-H bond in methane (Fig. 1).

Primers Targeting Propanotrophs and Butanotrophs

While PCR primers for the detection of monooxygenases, including methane monooxygenase found in methanotrophs, have been reported in literature (Stienkamp *et al.*, 2001) (Baldwin *et al.*, 2003), the use of PCR and primers in the detection of propane monooxygenase and butane monooxygenase is novel. As stated previously, although the detection of methanotrophs in a soil sample may indicate a source of subsurface methane gas, even a marked increase in presence cannot conclusively elucidate the gas' origins (Brisbane *et al.*, 1965). Conversely, natural subterranean sources of propane and butane derive exclusively from deposits of gas or crude oil. For this reason, PCR aimed at amplifying genes involved in propane and butane oxidation from a soil sample can serve as an accurate technique of discovering caches of crude petroleum.

In this thesis, three PCR primer sets were designed to specifically target propane and butane-oxidizing microorganisms. The primers were designed using conserved nucleotide sequences in genes encoding the large hydroxylase subunit of propane monooxygenases, *prm*A, and the gene encoding the alpha hydroxylase subunit of butane monooxygenases, *bmo*X, of several HCO species.

CHAPTER 2

MATERIALS AND METHODS

Primer Design

Gene sequences of the large hydroxylase subunit of propane monooxygenases (*prm*A) and the alpha subunit of butane monooxygenase hydroxylases (*bmo*X) are the basis of our primer designs. The *bmo*X, *prm*A, and *mmo*X nucleotide sequences in Table 1 were retrieved from the National Center for Biotechnology Institute (NCBI) nucleotide database in a query conducted on March 2, 2007. They were downloaded as FASTA files and aligned by the ClustalW protocol in Lasergene Megalign (DNASTAR, Madison, Wisconsin). Sections of the consensus sequence that exhibited the greatest sequence similarity were identified for possible primer annealing sites.

Lasergene PrimerSelect was used to design the sequence and length of candidate primers, based on the following criteria: similarity of melting temperatures (T_m), % GC content, and the likelihood of hairpin, primer dimer and self-primer dimer formation. Oligonucleotides were synthesized by Sigma Life Science (Sigma Aldrich Corporation, The Woodlands, TX) and suspended in PCR-grade water to create a 100uM stock that was stored at -20°C.

Table 1. Organisms and genes used to create the primers.

Organism Name	*GenBank Accession No.*	*Gene Name*	*Alkane Oxidized*
Positive Controls			
Gordonia sp. TY-5	AB112920.1	*prmA*	Propane
Mycobacterium sp. TY-6	AB250938.1	*prmA*	Propane
Pseudonocardia sp. TY-7	AB250941.1	*prm1A*	Propane
Pseudonocardia sp. TY-7	AB250942.1	*prm2A*	Propane
Methylibium petroleiphilum	CP000555.1	*prmA*	Propane
Rhodococcus jostii RHA1	NC008268.1	*prmA*	Propane
Methylibium petroleiphilum PM1	NC008825.1	*prmA*	Propane
Pseudomonas butanovora	AY093933	*bmoX*	Butane
Brachymonas petroleovorans	AY438629.1	*bmoX*	Butane
Negative Controls			
Methylosinus sporium SD56	AJ458525.1	*mmoX*	Methane
Methylosinus trichosporium	AJ458524.1	*mmoX*	Methane
Methylocella silvestris	CP001280.1	*mmoX*	Methane
Methylomonas LC1	DQ119051.1	*mmoX*	Methane
Methylocystis LR1	AJ458522.1	*mmoX*	Methane

Collection of Microbial Samples and DNA Isolation

Several microorganisms possessing a propane monooxygenase (PMO), butane monooxygenase (BMO), or multicomponent monooxygenase protein similar in sequence structure to PMO and/or BMO were chosen to serve as controls. *Pseudomonas butanovora* (ATCC 43655) and *Methylibium petroleiphilum* (ATCC BAA-1232) were chosen to serve as positive controls as they possess BMO/*bmo*X and PMO/*prm*A respectively. These two species were also included in the original CLUSTALW alignment used to design the primers. *Pseudomonas stutzeri (*ATCC 17588) and *Methylosinus trichosporium* (ATCC 49243) were chosen as negative controls as they possess toluene monooxygenase and methane monooxygenase respectively. The sMMO gene *mmo*X was selected as a negative control because of its amino acid sequence identity (64%) to butane monooxygenase (Sluis *et al.*, 2002). A freeze-dried culture of

each control organism was ordered from The American Type Culture Collection (ATCC) (Manassas, Virginia).

Approximately 0.1 g of each freeze-dried pellet obtained from ATCC was used to inoculate separate test tubes with 10 ml Tryptic Soy Broth (TSB). The cultures were then allowed to incubate at their appropriate temperatures, as indicated by the ATCC, until visible growth formed. The cultures were then streaked on Tryptic Soy Agar (TSA) plates and allowed to incubate at their optimum growth temperatures in gravity convection incubators until visible colonies formed. Colonies were used to inoculate new separate test tubes of TSB. Once visible growth formed, samples were centrifuged at 5,000 x g for 5 minutes at room temperature and the media was then decanted. Pellets were then placed in a 4°C refrigerator until DNA isolation.

DNA extraction was executed using PowerSoil® DNA Isolation Kits (MO BIO Laboratories, Carlsbad, CA). The entire bacterial cell pellet was transferred to a 2 ml PowerBead Tube and gently vortexed to mix. Solution C1 (60 µl) was added to the PowerBead Tube, inverted several times and briefly vortexed. The PowerBead Tube was placed in a Qbiogene Fast Prep Instrument (Carlsbad, CA) at 4.5 m/s for 30 seconds and then centrifuged at 10,000 x g for 30 seconds. Avoiding the pellet, supernatant was transferred to a clean microcentrifuge tube. Solution C2 (200 µl) was added to the extracted supernatant and vortexed for 5 seconds. The DNA extraction was allowed to incubate in a -20°C freezer for 15 minutes and then centrifuged for 1 minute at 10,000 x g. Supernatant (600 µl) was transferred to a clean microcentrifuge tube. Solution C3 (200 µl) was added. The microcentrifuge tube was vortexed briefly, incubated in a -20°C freezer for 15 minutes, and centrifuged for 1 minute at 10.0 x g. Again avoiding the

pellet, up to but no more than 750 µl of supernatant was transferred into a clean microcentrifuge tube. Solution C4 (1200 µl) was added to the supernatant and vortexed for 5 seconds. Approximately 675 µl of the supernatant was loaded onto a spin filter and centrifuged at 10,000 x g for 1 minute. The flow through was discarded and the remaining supernatant was added to the spin filter and centrifuged at 10,000 x g for 1 minute. Loading continued until all supernatant passed through the same filter. Solution C5 (500 µl) was added to the spin filter and centrifuged for 30 seconds at 10,000 x g. Flow through was discarded and the microcentrifuge tube spin filter was centrifuged again for 1 minute at 10,000 x g. The spin filter was placed into a new clean tube microcentrifuge tube, 100 µl of Solution C6 or PCR water was added to the center of the white filter membrane and was allowed to sit for 15 minutes. The centrifuge tube was centrifuged for 30 seconds at 10,000 x g. Completed DNA extractions were stored in a -20°C freezer.

Soil Samples and Isolates

Undergraduate student Brandon Winneker collected soil samples from the following sites: The Church of Jesus Christ (COJC) San Luis Obispo, CA (latitude 35.29419, longitude -120.674815), Santa Barbara Harbor (SBH) Santa Barbara, CA (34.407193, -119.693041), and the Plains Exploration & Production Company Arroyo Grande oil field (PXP) Arroyo Grande, CA (35.179281, -120.618811). Winneker isolated HCO organisms from these soils by plating them on carbon-source-free Bushnell-Haas media, then exposing them to propane or butane at 28°C. Under these conditions, growth on these media can only be propanotrophic or butanotrophic. The

isolates were then streaked on TSA for purity. HCO isolates derived from soil samples were used to provide propanotrophic and butanotrophic template DNA direct from the field. Frozen stocks of isolates *BC3*, *BS2*, *CPC2*, *CPP4* and *PP2* were made in sterile 20% glycerol. DNA was extracted from these soil isolates as described above.

DNA extractions were also made directly from soil sample COJC using a modified DNA extraction protocol from above. In place of cellular pellet, 1 g of each soil sample was loaded into a PowerBead Tube. A total of 4 g of soil was used. Replicates were combined after the Solution C4 step, when all supernatant was loaded onto the spin filter.

PCR Optimization

Optimization of the primers for PCR required the modification of various PCR protocol parameters: $MgCl_2$ concentration, Taq polymerase concentration, primer concentration, template DNA concentration, annealing temperatures, annealing times, extension times, and denaturing times. DNA was quantitated via BioPhotometer (Eppendorf, Hamburg, Germany). A Techne Touchgene Gradient Thermal Cycler (Techne Incorporated, Burlington, NJ) was used to elucidate optimal annealing temperatures, annealing temperature ranges, and times. PCR was performed in 50 µl reactions.

PCR product (7 µl) was analyzed on a 1% agarose-TBE gel, containing EtBr (1 mg • ml^{-1}), in 1.0X TBE buffer at 100V. The size of the PCR product was estimated by running 7 µl GeneRulerTM 1kb Plus DNA Ladder (Fermentas, Burlington, Ontario,

Canada) in an adjacent lane. Bands were visualized on a Bio-Rad (Hercules, CA) Gel Doc imaging system.

Detection Limits

To determine the sensitivity of the PCR assay, two separate detection limit experiments were performed. Dilutions of quantified template DNA (COJC, *M. petroleiphilum* or *Ps. butanovora*) were made in PCR-grade H_2O. The dilutions were incorporated into separate but identical PCR reactions during the PCR optimization period. PCR reactions were visualized on agarose gels to identify the lowest concentration of DNA capable of producing the expected 560-610 bp amplicon.

Detection Limits with DNA Interference

The second experiment was designed to elucidate the detection limit of the 955F25/1517R22 primer set under the influence of competitive interference by soil-derived DNA (negative soil). Three dilution sets served as the DNA volume for PCR reactions in this experiment. Dilution sets 1 and 2 contained positive control *M. petroleiphilum* DNA (3 ng • μl^{-1}) diluted with negative soil DNA (3 ng • μl^{-1}) and PCR-grade H_2O, respectively. Negative soil DNA was extracted from a soil core sampled from Wood River, Illinois. Dilution set 3 contained PCR-grade H_2O, with negative soil DNA as the diluent. The purpose of this last dilution set was to confirm negative soil DNA could not produce bands on its own. Identical volumes of each dilution were used

in separate reactions using the optimized PCR protocol. PCR product was visualized on agarose gels as previously described.

Gel Purification

When recovering PCR product from agarose gels, a Zymoclean™ Gel DNA Recovery Kit was used. The DNA fragment was excised from the from the 1% TBE agarose gel using a scalpel and transferred to a 1.5 ml microcentrifuge tube. Three volumes of Agarose Dissolving Buffer (ADB) was added to each volume of agarose excised from the gel (e.g. for a 100 mg agarose gel slice 300 µl of ADB was added). Tubes were incubated at 37°C for 10 minutes until the gel slice was completely dissolved. The melted agarose solution was pipetted into a Zymo-Spin™ Column in a collection tube, centrifuged at 10,000 x g for 60 seconds, and the flow-through discarded. Wash Buffer (200 µl) was added to each column and centrifuged at 10,000 x g for 30 seconds and the flow-through discarded. The wash step was repeated. The column was placed into a new 1.5 ml microcentrifuge tube and 10 µl PCR water was applied directly to the column matrix. Tubes were then centrifuged at 10,000 x g for 60 seconds to elute DNA.

Cloning and Sequencing

Using a PCR Ultra-Clean kit (MO BIO, Carlsbad, CA), five replicates PCR reactions were combined to help produce a successful ligation reaction. Five volumes of SpinBind solution were added to each reaction and repeatedly pipetted up and down in

order to mix the contents. The solution was then transferred to a spin filter unit within a 2 ml microcentrifuge tube. The microcentrifuge tube was centrifuged for 30 seconds at 10,000 x g and the filtrate discarded. Centrifugation and the discarding of the filtrate were repeated until all of the PCR-SpinBind solution was filtered, at which point all five PCR products were combined. SpinClean (300 µL) buffer was added to spin filter and centrifuged for 30 seconds at 10,000 x g. Eluate was discarded. Tubes were centrifuged again 120 seconds at 10,000 x g to remove any remaining fluid. The spin filter was transferred to a clean 2.0 ml collection tube. PCR water (60 µl) was added to the spin filter and allowed to incubate for 15 minutes. The microcentrifuge tube and spin filter were subjected to another round of centrifugation at 10,000 x g for 60 seconds and stored at –20°C.

For sequencing, amplicons were ligated into plasmids and transformed into *E. coli* for replication. Ligation was accomplished using the TOPO TA Cloning Kit (with pCR® 2.1-TOPO vector), (Invitrogen, Carlsbad, CA). Reactions (6 µl) were set up in 0.5 ml microcentrifuge tubes with the following reagents: 4 µl PCR product, 1 µl salt solution, and 1 µl pCR2.1 vector. Each reaction was incubated for 30 minutes at room temperature.

DNA transformations were conducted with One Shot® TOP10 Chemically Competent kits (Invitrogen, Carlsbad, CA). A water bath was brought to 42°C exactly and Lysogeny Broth-Ampicillin-XGal plates warmed to 37°C. SOC was defrosted at room temperature. One vial of One-Shot cells per sample was defrosted on ice. Each ligation reaction (2 µl) was added to a vial of cells TOP10 cells, swirled to mix, and incubated on ice. After 30 minutes cells were heat-shocked in the 42°C water bath for 30

seconds and immediately transferred back to ice. SOC (250 µl) was pipetted into each reaction. Tubes were then placed in a horizontal shaking incubator at 200 RPM, 37°C, for 1 hour. Cells were then plated on LB-Amp-XGal plates in 10 µl, 25 µl, and 50 µl aliquots using glass beads. Plates were incubated at 37°C for 18 hours. White colonies were picked and inoculated with 2 ml TSB-Amp at 37°C, 200 RPM, for 18 hours.

Plasmid Preparation

Zyppy Plasmid Miniprep Kit (Zymogen Research, Orange, CA) was used to isolate plasmid DNA from cells for DNA sequencing. In order to lyse the transfected cells, 100 µl of 7X Lysis Buffer was added directly to 600 µl of cell culture from the previous DNA transformation and mixed by inverting the microcentrifuge tube several times. Cold Neutralization Buffer (350 µl) was then added and mixed by inversion of the tube for complete neutralization. Tubes were centrifuged at 10,000 x g to pellet the cell debris. Avoiding the cell debris pellet, the resulting supernatant was transferred to a Zymo-Spin IIN column, placed into a collection tube and centrifuged at 10,000 x g for 15 seconds. Filtrate was discarded and the Zymo-Spin IIN column returned to the Collection Tube. Endo-Wash Buffer (200 µl) was added to the column and centrifuged at 10,000 x g for 15 seconds. Zyppy Wash Buffer (400 µl) was added to the column and centrifuged for 30 seconds at 10,000 x g. The Zymo-Spin IIN column was transferred to a new 1.5 ml microcentrifuge tube and then had 30 µl of Zyppy Elution Buffer added directly to the column matrix. The column was allowed to incubate at room temperature for 15 minutes to allow the plasmid DNA to elute from column and into the Elution

Buffer. Column and microcentrifuge tube were then centrifuged together for 15 seconds at 10,000 x g.

M13 PCR

PCR with the M13 primer set was used to determine the successfulness of the ligation and transformation steps. Vector isolated from clones (2 µl) was used in 25 µl M13 PCR reactions. Each reaction contained: 5 µl 5X Colorless GoTaq Flexi Buffer (Promega, Madison, WI), 2 µl dNTPs (10 mM), 2 µl MgCl2, 1 µl M13F, 1 µl M13R, 11.8 µl H2O, and 0.2 µl GoTaq DNA Polymerase (Promega, Madison, WI). PCR (7 µl) product was run on 1% agarose gels in TBE buffer at 100V for 45 minutes. GeneRulerTM 1kb Plus DNA Ladder was used to determine the approximate size of the product. Bands were visualized on a Bio-Rad (Hercules, CA) Gel Doc imaging system. Clones that produced a predicted 560-610 bp amplicon were used for sequencing.

Amplicon Sequencing

A combination of transformed cells and isolated plasmid were sent to Sequetech (Mountain View, CA), a DNA sequencing service, for PCR product sequencing. M13 primers were used in BigDye Terminator sequencing reactions. Rolling Circle Amplification (RCA) was applied to cell samples. RCA is an in vitro process whereby circular DNA can be amplified to produce high quality sequencing templates. DNA synthesis from circular DNA produces single-strand linear concatenated copies of the circular sequence. For samples producing weak and noisy signals, a proprietary

procedure known as BDX chemistry (Sequetech, Mountain View, CA) was employed. BDX chemistry is effective at getting through hard stops that result from particular sequence motifs including hairpins, stem/loops, or triple helices and high GC content sequences.

Sequence Editing and Phylogenetic Tree Building

Sequencing data returned from Sequetech were opened in the sequence-editing program Lasergene EditSeq (DNASTAR, Madison, WI). Sequence searches were conducted for the sequences of primers 955F25_Both and 1517R22 in both the forward and reverse direction. The Reverse Complement feature was utilized to format all sequences in the same direction. Detection of the correct primer sequences and amplicon length were used to confirm the presence of a target insert.

Alignments and phylogenetic trees of the sequencing data were needed to deduce the identities of all PCR products. To remove any primer bias in the sequence data, the primer sequences and vector sequences were removed from each raw sequence files before alignment. Contigs were built from the edited sequences with Lasergene SeqMan (DNASTAR, Madison, WI). These contigs were loaded into Lasergene MegAlign with the pared reference sequences of propanotrophs, butanotrophs, methanotrophs, ammonia-oxidizing bacteria, and tetrahydrofuran-oxidizing bacteria. Ammonia oxidizer *Rhodococcus rhodocrous amo*C (D37875.1) and tetrahydrofuran-oxidizer *Pseudonocardia K1 thm*A (AJ296087.1) share high amino acid sequence similarity with *prm*A and *bmo*X (Kotani *et al*., 2003). Along with the reference sequences, the edited sequences of

amplicons from soil sample COJC, propane isolates *CPC2* and *CPP4*, and positive controls *M. petroleiphilum* and *Ps. butanovora* were also included in the alignment. A ClustalW alignment was made with these sequences and a phylogenetic tree was constructed from the ClustalW alignment.

CHAPTER 3

RESULTS

Primer Design

Two regions of highly conserved nucleotide sequence were observed (Fig. 3) in the alignment that included nine *prm*A and *bmo*X reference sequences (Table 1). Sections 931-955 and 1517-1538 of the alignment consensus sequence were chosen as the basis of forward and reverse primer design. In a similar alignment, methane monooxygenase *mmo*X sequences were also included (Table 1). The shorter *mmo*X sequences aligned at section 931-955 but did not possess the 1517-1538 section of the consensus sequence. This fact contributed to the selection of these sections as primer annealing sites. Sections 931-955 and 1517-1538 also possessed stretches of exact similarity that allow for strong specific annealing of the 3' end of their respective primers to their respective annealing targets. Using the primer sets and the sequences from which they were based, 560-610 bp PCR products were predicted. The section of the consensus sequence that is flanked by the two selected primer sites (956-1516) possesses an additional site of high sequence agreement (1133-1166) (Fig. 4). Its position within future PCR products could serve as a promising qPCR TaqMan probe-annealing site.

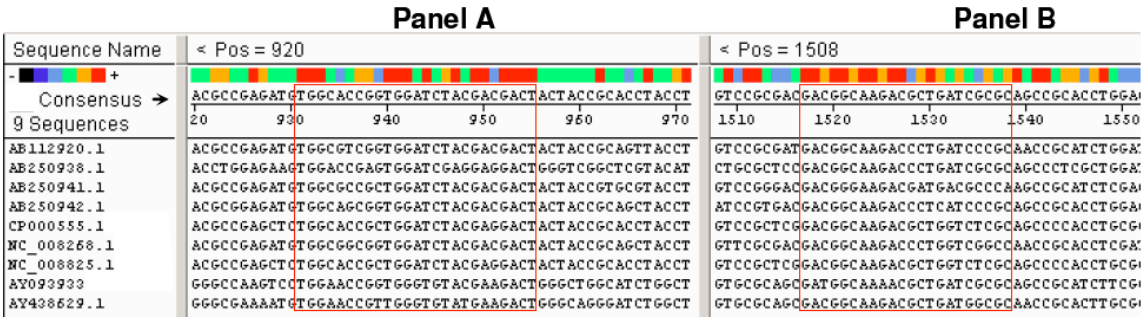

Figure 3. Portions of a DNA sequence alignment of the hydroxylase subunit genes from the sources listed in Table 1. Panel A is the portion of the alignment used to design primer 955F25_both. Panel B is the portion of the alignment used to design primer 1517R22. Red boxes surround relevant primer sequences. The left-hand column identifies each reference sequence used by its Gen Bank Accession number as detailed in Table 1. The last two sequences are from *bmo*X genes.

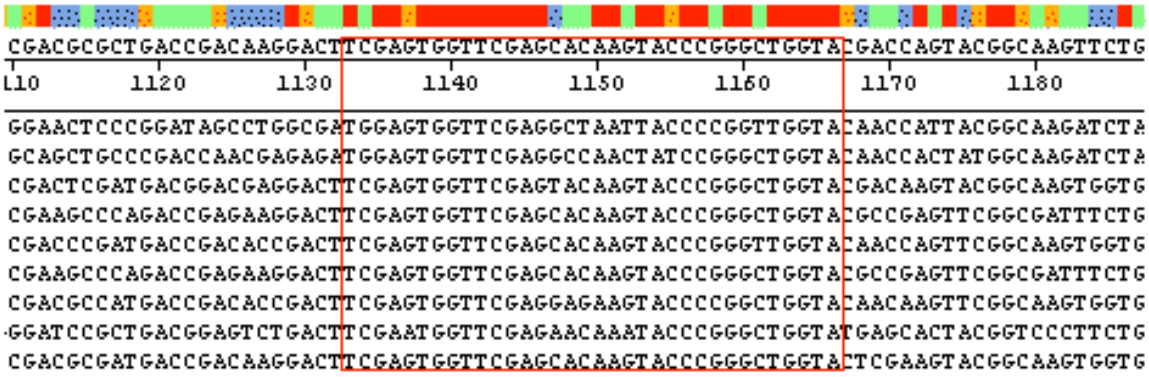

Figure 4. Portion of the ClustalW *prm*A/*bmo*X alignment consensus sequence (1133-1166) fit for a potential qPCR TaqMan probe-annealing site.

Three primer sets were proposed, each utilizing a different forward primer with the same 1517R22 reverse primer (Table 2). The 955F25_both primer was designed to detect both propane and butane oxidizers, 955F25_propane was designed to detect propane oxidizers and 955F25_butane was designed to detect butane oxidizers. Sequence divergences necessitated a degenerate design for the primers to capture a wide range of both propane-oxidizing and butane-oxidizing microorganisms. Degeneracy, or doping,

was deployed in ratios that reflected ratios in the alignment. An inosine was also employed to bind to various nucleotides.

Table 2. Primer nucleotide sequences.

Primer	*Sequence*
955F25_both	TGGCACCGGTGG(A:G 4:1)T(C:G 4:1)TACGAIGACT
955F25_propane	TGGCACCGGTGGATCTACA(C:G 4:3)GACT
955F25_butane	TGGCACCGGTGGGTGTACGAAGACT
1517R22	GCGCGATCAG(C:G 3:2)GTCTT(G:C 9:1)CC(G:A 1:9)TC

[A] Ratios denote the specific ratio of specific nucleotides at that position in the oligonucleotide.

[B] Iosines, which can bind to A, C or T, are represented by *I*.

PCR Optimization

The three sets of primers displayed varying levels of success. None of the primer sets produced bands from negative controls *M. trichosporium* and *Ps. stutzeri*. The primer set that included 955F25_propane, was successful in generating a PCR product in the presence of *M. petroleiphilum* and the COJC soil sample at annealing temperatures ranging from 66-68 °C and $MgCl_2$ concentrations between 5-8 mM (Table 3). The primer set that included 955F25_butane was only successful in generating a product from *Ps. butanovora* but not soil at annealing temperatures (66-69 °C) and $MgCl_2$ concentration between 5-8 mM. 955F25_butane also began producing a band with the negative control *M. petroleiphilum* at annealing temperatures below 65 °C. However, 955F25_both was able to produce PCR product from *M. petroleiphilum*, *Ps. butanovora*, and the COJC soil samples. Since the 955F25_both primer successfully

amplified both propane, butane monooxygenase genes, and COJC soil (Table 4), it was chosen to proceed with for amplicon cloning and sequencing.

Table 3. PCR optimization results for each primer set with various DNA templates with various PCR protocols

Primer Set	*DNA Template*	*Primer Annealing Temperature (°C)*	*[$MgCl_2$] (mM)*	*Proper Band (Yes/No)*
955F25_propane	*M. petroleiphilum*	66-69	5-8	Yes
	Ps. butanovora	64-69	5-8	No
	COJC Soil	66-68	6-8	Yes
955F25_butane	*M. petroleiphilum*	$\leq$ 66	5-8	Yes
	Ps. butanovora	66-69	5-8	Yes
	COJC Soil	64-69	5-8	No
955F25_both	*M. petroleiphilum*	65-69.7	6-8	Yes
	Ps. butanovora	65-69	6-8	Yes
	COJC Soil	65-68	6-8	Yes

Table 4. PCR product sizes using the 955F25_both primer set.

DNA Template	*Band Size (bp)*
M. petroleiphilum	604
Ps. Butanovora	580
COJC	580, 601, 604, 613

* *COJC* – Church of Jesus Christ soil sample

After hundreds of PCR reactions, the optimized 955F25_both/1517R22 protocol called for 50 µl reactions of: 1µl genomic DNA, 1.5 µl 955F25_both (10 µM), 1.5 µl 1517R22 (10 µM) (Sigma Aldrich Corporation), 5 µl 10X PCR Gold Buffer (Applied Biosystems Foster City, CA), 3 µl dNTPs (10 mM), 2 µl BSA (20 µg • ml^{-1}), 16 µl $MgCl_2$

(25 mM), 0.3 µl Taq Gold (5 U • μl^{-1}), and 19.7 µl PCR water to bring the total volume up to 50 µl. PCR was performed in an Applied Biosystem 96-well Thermal Cycler (Foster City, CA). The first stage was a 94°C denaturation for 10 minutes. This was followed by 35 cycles of the following parameters: denaturing for 1 minute at 94°C, annealing for 1 minute at 66°C., and extension for 2 minutes at 72°C. After the final cycle, an extra extension stage lasted for 10 minutes before the reactions were held at 4°C.

During the optimization period, PCR smearing began to appear in the agarose gel images (Fig. 5A) between 3/18/09 - 8/5/09. Smearing ran the complete length of the gel (approximately 0.05-20 kb) with soil sample COJC and positive controls as template. Also, during the optimization period, bands disappeared and reappeared in agarose gel images over the span of hundreds of PCR reactions (Fig. 5B) between 3/18/09 - 8/5/09. PCR protocols that had previously produced a proper-sized band, failed in subsequent attempts.

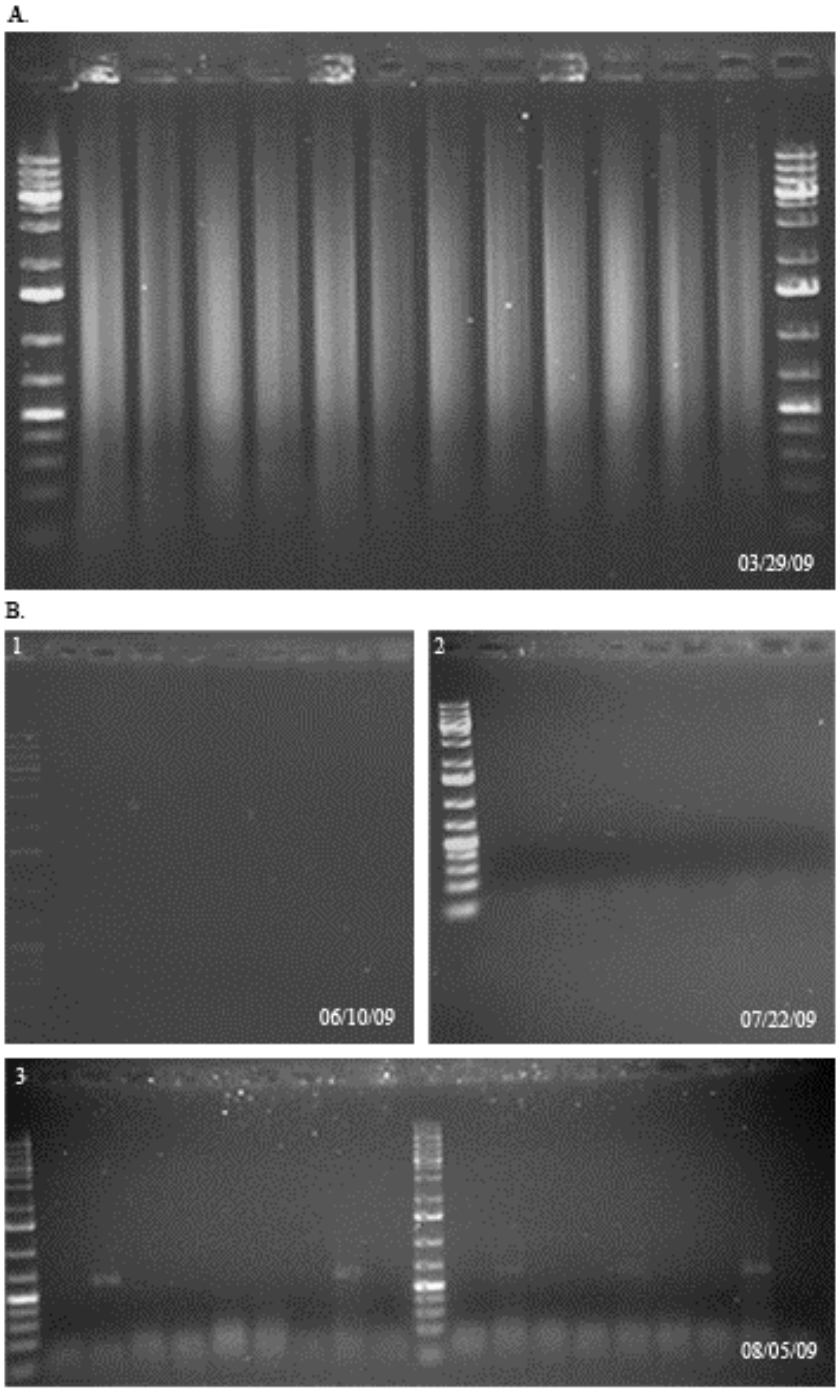

Figure 5. Problems with PCR. Each gel contains 7 µl GeneRulerTM 1kb Plus DNA Ladder. Lanes contain PCR reactions with template DNA from COJC soil, *M. petroleiphilum*, and *Ps. butanovora*. (A) Experimental lanes exhibited PCR smearing. Image captured 3/29/09. (B) Example gel images demonstrating the disappearance of PCR (B1) product and its reappearance (B3). Bands were undetectable over hundreds of PCR reactions from 03/18/09 until 08/05/09 when positive control *M. petroleiphilum* reappears. Images were captured: B1) 06/10/09, B2) 07/22/09, and B3) 08/05/09.

HCO Isolate PCR

Brandon Winneker characterized each HCO soil isolate (Table 5). HCO isolates CPC2 and CPP4 each produced a single band (610 bp) of target-sized (560-610 bp) PCR product while soil isolates *BC3*, *BS2* and *PP2* all produced either non-target-sized or multiple bands (Table 5). Attempts to recover target-sized bands from *BC3*, *BS2*, and *PP2* were made via gel purification kits. Isolate clone *BC3* returned shorter and longer sequences (434 bp and 661 bp) and isolate clone *BS2* produced a 552 bp product with both primer sequences included. *PP2* produced short amplicons (281-472 bp) with both primer sequences included and long products (742 bp) with the 1517R22 primer missing.

Table 5. Characterized soil isolates that produced PCR product and whether a contig could be constructed from the sequencing data.

Isolate	*Gram Stain*	*Cell Morphology*	*Catalase*	*Oxidase*	*Amplicon Sizes (bp)*	*Contig*
BC3	Negative	Rod	-	+	434, 661	-
BS2	Negative	Coccobacilli	+	-	552	-
CPC2	Negative	Rod	-	-	601	+
CPP4	Negative	Rod	+	-	601	+
PP2	Negative	Rod	Delayed +	-	281-472, 742	-

Detection Limit Experiment

For the detection limit experiments, an average bacterial genome size of 5 Mbp and a conversion factor of 978 Mbp • pg DNA^{-1} were used. It was also assumed in this experiment, that cells possess only one genome copy per cell. All reactions containing *M. petroleiphilum* or *Ps. butanovora* produced target-sized bands down to a concentration of 10 pg (2×10^3 target gene copies) template DNA per PCR reaction.

Reactions containing 1 pg (1.96 x 10^2 target genes copies) DNA per PCR reaction failed to produce detectable bands. DNA isolated from COJC soil, produced target-sized bands down to 200 pg template DNA per reaction.

DNA Interference Detection Limit Experiment

Dilution Set 1 produced a single band down to 3 pg DNA (5.88 x 10^2 gene copies) • PCR reaction^{-1} (Table 6), while Dilution Set 2 produced a single band down to 30 pg (5.88 x 10^3 gene copies). No secondary bands were observed. The control, Dilution Set 3, did not produce any bands. All bands produced were within target product size range (580-613 bp).

Table 6. DNA interference detection limit results displaying the pg and number of gene copies of DNA in each PCR reaction and whether or not a band was produced. All bands produced were within target product size range (580-613 bp).

(Dilution Set No.)	*300 pg* DNA *(6 x 10^4 genes)*	*30 pg* DNA *(6 x 10^3 genes)*	*3 pg* DNA *(6 x 10^2 genes)*	*1 pg* DNA *(2 x 10^2 genes)*
(1) *M. petroleiphilum* in negative soil DNA	+[a]	+	+	-[b]
(2) *M. petroleiphilum* in PCR-Grade H_2O	+	+	-	-
(3) PCR-Grade H_2O in negative soil DNA	-	-	-	-

[a] +, reaction produced PCR product

[b] -, reaction did not produce PCR product

Confirming Amplification: PCR Product Sequencing

To confirm the accuracy of PCR amplification, sequencing of PCR products is required to identify the template DNA that was amplified by a primer set. This provides an indication of a primer set's specificity. Sequence identification was accomplished by entering sequencing data into NCBI BLAST queries, by using the sequencing data to construct phylogenetic trees, or both.

Amplicons resulting from the 955F25_both/1517R22 primer set were sequenced. Sequencing data can be found in the Appendix. The amplicon from *M. petroleiphilum* was identified by NCBI BLASTn Megablast as a portion of the large hydroxylase subunit of propane monooxygenase of *Methylibium petroleiphilum* (CP000555.1) score = 1029, Identities = 557/557 (100%), Gaps = 0/557, and E value = 0.0. The amplicon produced from *Ps. butanovora* was identified by NCBI BLASTn Megablast as a portion of butane monooxygenase hydroxylase (BMOH) alpha subunit of *Pseudomonas butanovora* (AY093933.3) (Score = 979 bits (530), E value = 0.0, Identities = 532/533 (99%), Gaps = 0/533 (0%)). NCBI BLASTn Megablast largely identified COJC soil-derived amplicons as portions of methane monooxygenase-like, methane monooxygenase, or putative monooxygenase genes. Propanotrophic isolates *CPC2-B* and *CPP4-G* were found to be almost exact matches to soil amplicons *COJC17*and *COJC38*, and extremely similar to *COJC14* (Identities = 94%).

Phylogenetic Tree Building

A phylogenetic tree was assembled from a ClustalW alignment of the sequences (Fig. 6). In preparation for the alignment, all experimental sequencing data and reference sequences were pared down to include only the regions between and not including primer-binding sites. Alignment sequences included in Table 1 and experimental sequencing data included soil-derived amplicons (COJC), *M. petroleiphilum*-derived amplicons, *Ps. butanovora*-derived amplicons, HCO isolates (*CPC2*, *CPP4*), and negative control genes *mmo*X*, amo*C*,* and *thm*A.

The largest clade within the phylogenetic tree was composed of soil-derived COJC amplicons. The few exceptions non-soil-amplified sequences within this clade were from propane-oxidizing isolates *CPC2*-B and *CPP4*-G, which were identical to soil amplicons *COJC17* and *COJC38*, and closely related to *COJC14*. This large clade of COJC amplicons is then most closely related to a clade of *M. petroleiphilum* sequences and a clade of propanotrophic reference sequences. The *M. petroleiphilum* (PCR) amplicons share the exact same sequence as the *M. petroleiphilum* reference sequences. Soil amplicon *COJC28* is also included within the *M. petroleiphilum* clade. The propanotrophic reference clade contains soil amplicon COJC10 and a subclade of COJC23 and *Pseudonocardia TY-5 prm1A* (AB250941.1).

Another major clade was composed of butanotrophic and methanotrophic sequences. One subclade included *Ps. butanovora* (PCR) amplicons, *Ps. butanovora* (AY093933) and *B. petroleovorans* (AY438629) while methanotrophic *mmo*X reference sequences constituted the second subclade. *Rhodococcus rhodochrous amo*C was also

placed in the butanotrophs and methanotrophs clade. *Pseudonocardia K1 thm*A was placed furthest away from the other sequences.

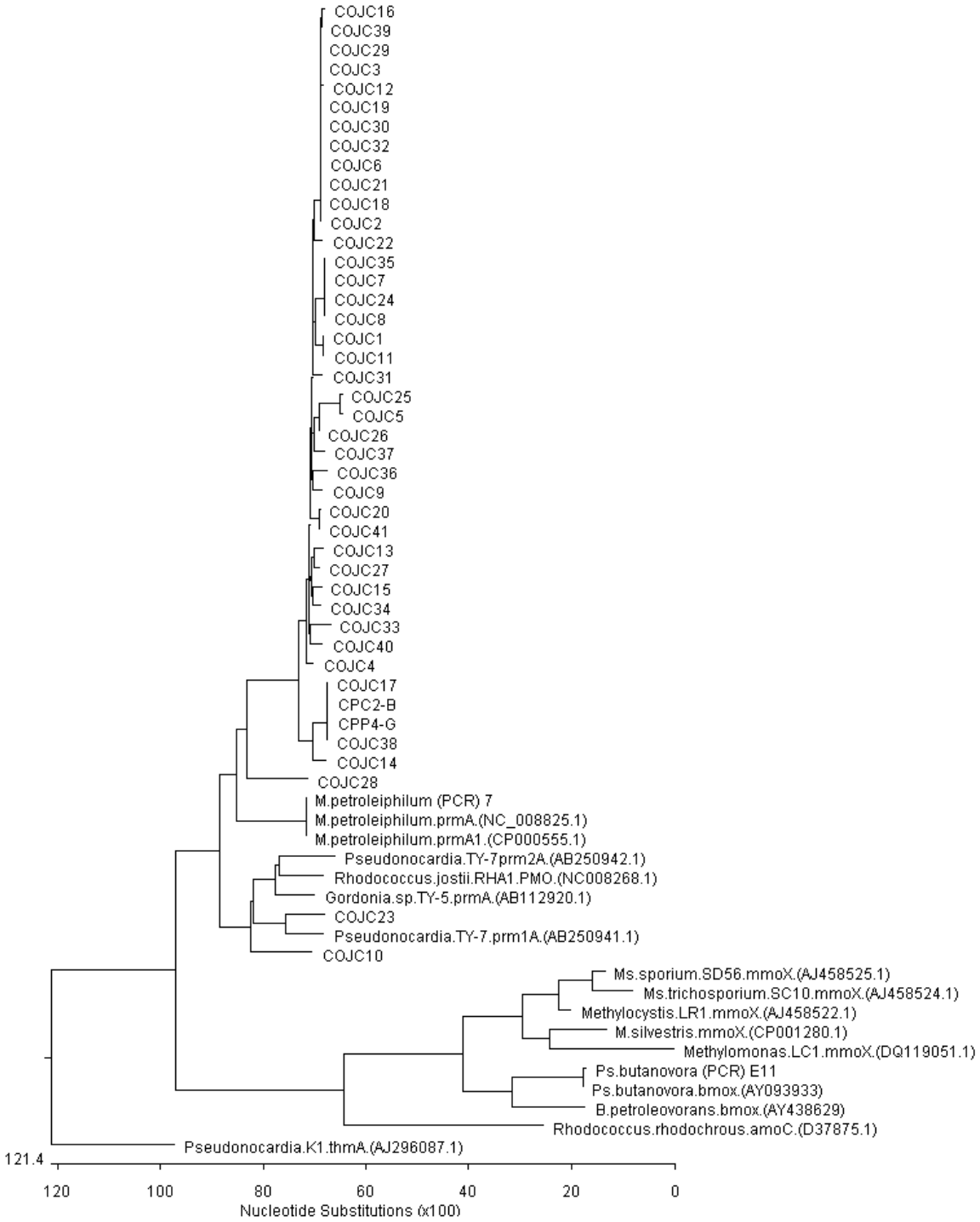

Figure 6. Phylogenetic tree – ClustalW alignment includes sequenced PCR products from: soil sample DNA (*COJC*), *Methylibium petroleiphilum* (PCR), *Pseudomonas butanovora* (PCR), and soil isolates *CPC2*-B and *CPP4*-G. Alignment contains Genbank reference sequences: Table 1 organisms, *mmo*X organisms *Methylosinus trichosporium*, *Methylocystis sp. LR1*, *Methylosinus trichosporium SC10*, *Methylocella silvestris* and *Methylomonas LC1*, *Pseudonocardia sp.K1* tetrahydrofuran monooxygenase α subunit (*thm*A), *Rhodococcus rhodochrous* (*Gordonia rubripertincta*) ammonia monooxygenase epoxidase subunit (*amo*C).

CHAPTER 4
DISCUSSION

The intent of this study was to create a primer set capable of detecting propane and butane-oxidizing bacteria through PCR. Although not completely refined, the primer sets accomplish the major goals set out at the beginning of the experiment.

Primer Specificity with Control DNA

The purpose of this study was to develop a PCR assay to rapidly and accurately detect propane-oxidizing and butane-oxidizing microorganisms. Through PCR with positive controls *M. petroleiphilum* and *Ps. butanovora*, it was clearly shown that the primer pair 955F25_both and 1517R22 is capable of specifically amplifying regions of the hydroxylase subunits of both propane monooxygenase and butane monooxygenase. *M. petroleiphilum* and *Ps. butanovora* produced a 604 bp and a 580 bp amplicon respectively. The primers failed to produce any bands in the presence of negative control *M. trichosporium* methane monooxygenase and negative control *Ps. stutzeri* toluene monooxygenase template DNA.

PCR Issues

PCR smearing plagued the experiment for months and is difficult to account for. While occurrences were irregular, the problem is well documented by many PCR

laboratories. Some attribute smearing problems to the gradual build-up "amplifiable DNA contaminants" specific to particular primers or to the prolonged and repeated use of old primers (Han *et al.*, 2006). Others have attributed smearing to inappropriate thermal cycling temperatures, reagent concentrations or DNA concentrations. As protocols remained the same and reagents were replaced regularly, the cause of smearing may lie in the contamination or degradation of genomic DNA over time as PCR relies on pure, unfragmented DNA templates (Burgman *et al.*, 2001).

The disappearance and reappearance of bands over time is also difficult to explain. While protocols remained the same and the reagents were replaced regularly, the erratic amplification appeared to be DNA related. Positive control DNA continued to produce normal bands semi-regularly while DNA extracted from soil failed on a more consistent basis. When additional DNA was extracted from soil, fresh extractions temporarily led to the recovering of bands but the effectiveness of new DNA extractions diminished over time. The repeated freezing and thawing of unprocessed samples over a long period of time can cause DNA to degrade (Ross *et al.*, 1990).

Primer Specificity with Heterogeneous Soil Samples and Soil Isolates

While the ability to produce proper-sized PCR products from soil is evident, the identities of these amplicons remain less than certain. Both methods utilize sequence alignments to compare sequences in order to elucidate sequence identities. NCBI BLAST queries compare query sequences with a database of identified sequences and returns scored matches above a set threshold. In addition to providing sequence matches,

NCBI BLAST results can provide the organism and area of genome where the matched sequence originated. Phylogenetic trees compare multiple input sequences to each other to infer sequence relationship and identification. Diagrammatically, similar sequences are grouped closer together than more dissimilar, distantly related sequences.

Several sets of amplicons were sequenced in this experiment. Firstly, amplicons from positive controls *M. petroleiphilum* and *Ps. butanovora* were sequenced. Each primer's design was partially based on reference sequences of both positive controls, so each primer set should accurately amplify the targeted DNA of the positive controls. A level of accuracy can be established if positive control sequencing data and positive control reference sequences match identically. The next set of amplicons to be sequenced was one that was produced from soil sample COJC. DNA extracted from soil samples provides a mixture of HCO DNA and background non-target DNA template that cannot be replicated in the lab. Sequencing data from these amplicons provided information on the specificity and diversity captured from environmental samples. From these samples, we can also recognize whether non-target DNA is being amplified and identify the assortment of HCO organisms that are being captured by the primer set. The propane and butane isolates offer an opportunity to PCR amplify from confirmed butane or propane oxidizers isolated from environmental samples.

In the constructed tree, all COJC soil amplicon sequences displayed closest sequence similarity to propane monooxygenase sequences: the Genbank reference sequences, soil isolate sequences and sequence from the positive control *M. petroleiphilum*. Soil isolate *COJC23* was grouped exclusively within a subclade with the propane monooxygenase from *Pseudonocardia sp. TY-7 prm1A* (AB250941.1). *COJC10*

was also grouped with the propane monooxygenase reference sequences. Soil amplicon sequences *COJC17* and *COJC38* were identical to sequences from propane-oxidizing soil isolates *CPC2*-B and *CPP4*-G. *CPC2*-B and *CPP4*-G are derived from confirmed propanotrophic organisms. These facts strongly suggest that the COJC soil-derived amplicons, especially *COJC17* and *COJC38*, are also derived from propanotrophic microbes. COJC soil isolates were grouped most closely with propane monooxygenase reference sequences while the butane monooxygenase reference sequences were grouped more closely with the methane monooxygenase reference sequences. While NCBI BLASTn Megablast results identified COJC isolates as methane-monooxygenase-like, it is clear from the ClustalW alignment, that they are in fact more closely related to propane monooxygenases than methane monooxygenases. Together, these findings show that the 955F25_both/1517R22 primer set amplified propane monooxygenase genes from soil. The COJC soil amplicons in these experiments are most likely from yet to be identified propane monooxygenase genes.

Some isolates failed to produce any useable sequence despite the use of a proper sized insert within each vector. *BC3*, *BS2*, and *PP2* all failed to produce full-length amplicon sequences capable of contig building. None of these sequence data contained a forward and or reverse primer sequence within them, which may indicate a PCR failure rather than a sequencing failure. Sequencing failures are often caused by a high GC%, causing hard stops, and degraded amplicon sample, resulting in short interrupted DNA template. Short amplicons are difficult to sequence, and could be the cause of the nonsense sequence data.

The disparity between propanotrophic and butanotrophic clones recovered from soil sample COJC may have been caused by either the absence of butanotrophic template in the original COJC soil sample or the propanotrophic bias of the 955F25_both/1517R22 primer set design. Degeneracies within the 955F25_both/1517R22 primer set may be the source of some of its past and future shortcomings. Both degenerate sites in the 955F25_both primer are doped favorably in ratios (4:1) that favor propanotrophic template binding. None of the degenerate sites in the reverse primer are biased toward either propanotrophs or butanotrophs. The use of this primer set involved the deployment of four unique forward primers and eight unique reverse primers in every PCR reaction. And to different extents, all were capable of dimerization, self-dimerization, concatemerization and perhaps most significantly, indiscriminate annealing and amplification. Less discriminate annealing and amplification could have lead to the production of a false positive result: the identification of a HCO population where none exists. Further discussion of primer sequence degeneracies can be found in the Recommendations chapter.

Detection Limits

Positive controls produced detectable bands down to 10 pg DNA (2×10^3 target gene copies)• 50 µl PCR reaction^{-1} and soil down to 200 pg DNA (4×10^4 total gene copies)• 50 µl PCR reaction^{-1} in the first detection limit experiment. The disparity between detection limits is due to HCO communities representing only a fraction of the organisms living in soil. According to the results of the first detection limit experiment,

we can approximate that roughly 2-10 pg of the 200 pg of the soil-derived DNA required for each PCR reaction (1-5%), was derived from HCO organisms. In line with our previous assumptions about average genome size, if 100 µl DNA was recovered from 4 g of COJC soil, we can estimate a soil cellular concentration of 5×10^5 cells•g soil^{-1}. Five percent ($4 \times 10^3 – 2 \times 10^4$ cells•g soil^{-1}) are HCO organisms. This number is certainly an underestimate, due to extraction inefficiency and extraction bias. About 7% of the yield is lost due to the MoBio Powersoil protocol and anywhere from 10-60% remains unrecovered depending on soil type (Sagova-Mareckova *et al*., 2008). Taking this into account, the COJC soil HCO population concentration was originally between 1.6×10^3 – 1.8×10^4 cells • g soil^{-1}. This estimate simultaneously classifies COJC soil as borderline anomalous and limits the assay to detecting only anomalously large communities ($\geq 10^4$ cells • g soil^{-1}) (Wagner *et al.*, 2002). To our advantage, this creates a plus/minus PCR assay; one that produces detectable PCR product from only anomalously high HCO soil populations. These are only limited estimates. More accurate calculations could be made from future soil-spiking experiments.

The DNA interference experiment reinforces the detection limit results from the first experiment. PCR reactions with Dilution Set 2, made up of a dilution of *M. petroleiphilum* in PCR-grade H_2O, set the estimate for detection limits between 3-30 pg • PCR reaction^{-1}. Unexpectedly, Dilution Set 3 displayed the lower detection limit (1-3 pg • PCR reaction^{-1}). If competitive interference had occurred, PCR-grade H_2O would have produced the stronger bands at the more dilute concentrations. DNA from the negative control sample did not interfere with producing PCR product and may have even bolstered it. Dilution Set 3 did no produce any bands on its own. However, the negative

soil DNA may have contained enough HCO organism DNA to increase the intensity of the bands produced by *M. petroleiphilum.* The negative control's non-target DNA did not interfere with *M. petroleiphilum* producing a target-sized product and did not produce secondary bands.

CHAPTER 5

RECOMMENDATIONS

Any future use of the 955F25_both/1517R22 primer set, should begin with the modification of primer sequences. Removing selected, or all, degeneracies built into these primers could increase annealing specificity and simultaneously reduce the rate of nonspecific primer binding, therefore reducing the rate of nonspecific amplification. In addition, fine-tuning of the degeneracy ratios may help capture a wider range of HCO organisms. As they stand, the designs of 955F25_both and 1517R22 bias them to anneal more preferentially to propane monooxygenase template DNA than butane monooxygenase. Modifying annealing temperatures may also help capture a winder range of HCO organisms by allowing some primers to anneal more and others less preferentially than at current annealing temperatures. Sequencing the soil sample amplicons of a different annealing temperature may reveal amplicons more closely related to butane monooxygenase. A viable alternative would be the substitution of the 955F25_both primer with the 955F25_propane and 955F25_butane primers - separately or in conjunction with one another.

Adaptation of the 955F25_both/1517R22 primers set to a qPCR TaqMan platform is the next logical phase of the project. Doing so, would increase the assay's specificity and add a quantitative facet to the assay. A TaqMan assay would allow for the detection and quantitation of a specific sequence in DNA samples. Section 1136 - 1162 is an excellent candidate for a TaqMan probe-binding site (Fig. 4). The TaqMan hydrolysis probe adds an additional layer of specificity by requiring template DNA to possess

section 1136-1162 in addition to sections 931-955 and 1517-1538 to produce the characteristic qPCR fluorescent signal. Additionally a qPCR platform would allow for the quantitation of DNA and by proxy, the quantitation of an HCO population in a soil sample.

Additional positive and negative controls would also help optimize the accuracy of these primers. Only two positive controls were shown to have produced target PCR product. More controls would further solidify the primer set's range of detection and its accuracy. Due to cost-related issues, attaining supplementary cultures from the ATCC proved expensive. Known butane and propane oxidizers would continue to serve well as positive controls and organisms with closely related monooxygenases as advantageous negative controls. Negative controls could also be identified by the sequences gained from future sequencing of soil sample amplicons because these organisms that are providing the false positive results.

APPENDIX

PCR PRODUCT SEQUENCES

1) COJC1 (3-04-10_COJC_A3)

GGAACGGACGAAACCCATGTCCTTGATCGCGTCGGCCAGATCCCAGCCGTGA
TACATGCTCTCCCACTCGCGATGACGCCGCTGAAGCGGCCCATCGCCGGCGT
CGGAAGCCCCTGATATTCGCCTTGGAAAGCGACCTTGTGGGTCCAGCGGTCG
ACTTCATGGCCGTAGGTGTAGATCTCGCCGTCCACCTCATCGACCACGATGTC
CTCGCGGATCAGGCAGGGAACGAGGTTGGACCAGCAGCGATGCGGATAGAC
GTAGCCGGTATCGACGAAGGTGATCGGCGGATTGCCGGGTTTCGACAGCTTG
GCATAGTTTTCCCACCAAATACCGAATTCATCGTACCAGCCGGGATACTTGTG
CTCGAACCACTCGAAGTCACGCTCGGTCATCGCCTCGATGCGCCAGAAGTTG
GCCCACCAGCCGGCGGAGAAGAACTGGGCGACTTTGTGGACGTAGTTCTTCT
TGACGATCCGGTCGAAAGCTTCGTGGACATCGTCGTGGTGGATCTTGATGCC
GTACTTCTCCAGCGGCAGCATGTAGGTCCGGTAAT

2) COJC2 (3-04-10_COJC_C1)

GGAACGGACGAAACCCATGTCCTTGATCGCGTCGGCCAGATCCCAGCCGTGA
TACATGCTTTCCCACTCGCGATGACCGCTGAAGCGACCCATCGCCGGGGTCG
GCCGGCCCTGATATTCGCCCTGGAAGGCGACCTTGTGGGTCCAACGATCCAC
TTCATGGCCGTAGGTGTAGATTTCGCCGTCCACCTCATCGACCACGATATCCT
CGCGGATCAGGCAGGGAACGAGGTTGGACCAGCAGCGATGCGGATAGACGT
AGCCGGTATCGACGAAGGTGATCGGCGGATTGCCGGGTTTCGACAGCTTGGC
ATAGTTTTCCCACCAAATACCGAATTCATCGTACCAGCCGGGATACTTGTGCT
CGAACCACTCGAAGTCACGCTCGGTCATCGCCTCGATGCGCCAGAAGTTGGC
CCACCAGCCGGCGGAGAAGAACTGCGCGACCTTGTGGACGTAGTTCTTCTTG
ACGATCCGGTCGAAGGCTTCGTGGACATCGTCGTGGTGGATCTTGATGCCGT
ATTTCTCCAGCGGCAGCATATAGGTCCGGTAAT

3) COJC3 (3-04-10_COJC_D2)

GGAACGGACGAAACCCATGTCCTTGATCGCGTCGGCCAGATCCCAGCCGTGA
TACATGCTTTCCCACTCGCGATGACCGCTGAAGCGACCCATCGCCGGGGTCG
GCCGGCCCTGATATTCGCCCTGGAAGGCGACCTTGTGGGTCCAACGGTCCAC
TTCATGGCCGTAGGTGTAGATTTCGCCGTCCACCTCATCGACCACGATATCCT
CGCGGATCAGGCAGGGAACGAGGTTGGACCAGCAGCGATGCGGATAGACGT
AGCCGGTATCGACGAAGGTGATCGGCGGATTGCCGGGTTTCGACAGCTTGGC
ATAGTTTTCCCACCAAATACCGAATTCATCGTACCAGCCGGGATACTTGTGCT
CGAACCACTCGAAGTCACGCTCGGTCATCGCCTCGATGCGCCAGAAGTTGGC
CCACCAGCCGGCGGAGAAGAACTGCGCGACCTTGTGGACGTAGTTCTTCTTG
ACGATCCGGTCGAAGGCTTCGTGGACATCGTCGTGGTGGATCTTGATGCCGT
ATTTCTCCAGCGGCAGCATATAGGTCCGGTAAT

4) COJC4 (3-04-10_COJC_D4)

GGAACGGACGAAACCCATGTCCTTGATCGCGTCGGCCAGATCCCAGCCGTGA
TACATGCTCTCCCACTCGCGATGACCGCTGAAGCGGCCCATCGCCGGGGTCG
GCCGGCCCTGGTATTCGCCCTGGAAGGCGACCTTGTGGGTCCAGCGGTCGAC
TTCATGGCCGTAGGTGTAGATCTCGCCGTCGACCTCATCGACCACGATGTCCT
CGCGGATCAGGCAGGGGACGAGGTTGGACCAGCAGCGGTGCGGATAGACGT
AGCCGGTATCGACGAAGGTGATCGGCGGGTTGCCGGGCTTCGACAGCTTGGC
ATAGTTTTCCCACCAGGCGCCGAATTCATCGTACCAGCCGGGATACTTGTGCT
CGGACCACTCGAAGTCACGCTCGGTCATTGCCTCGATGCGCCAGAAGTTGGC
CCACCAGCCAGCGGAGAAGAACTGAGCGACCTTATGGACGTAGTTCTTCTTG
ACGATCCGGTCGAACGCTTCGTGGACGTCGTCGTGGTGAATCTTGATGCCGTA
CTTCTCCAGCGGCAGCATGTAGGTCCGGTAAT

5) COJC5 (3-04-10_COJC_G4)

GGAACGGACGAAACCCATGTCCTTGATCGCGTCGGCCAGATCCCAGCCGTGA
TACATGCTCTCCCACTCGCGATGACCGCTGAAGCGGCCCATGGCGGGTGTCG
GGCGCCCCTGATATTCGCCCTGGAAAGCGACCTTGTGGGTCCAGCGATCGAC
TTCATGGCCGTAGGTGTAGATTTCGCCATCCACCTCATCGACCACGATGTCTT
CGCGGATCAGGCAGGGAACGAGGTTCGACCAGCAGCGGTGCGGATAGACGT
AGCCGGTATCGACAAAGGTAATCGGCGGGTTGCCGGGCTTCGACAGCTTGGC
ATAGTTTTCCCACCAGGTGCCGAATTCATCGTACCAGCCGGGATACTTGTGCT
CGAACCACTCGAAGTCACGCTCGGTCATCGCCTCGATGCGCCAGAAGTTGGC
CCACCAGCCGGCCGAGAAGAACTGTGCGACCTTGTGGACGTAGTTCTTCTTG
ACGATCCGATCGAAGGCTTCGTGGACGTCGTCGTGATGGATCTTGATGCCGT
ATTTCTCCAGCGGCGGCATATAGGCCCGGTAAT

6) COJC6 (3-11-10_COJC_A3)

GGAACGGACGAAACCCATGTCCTTGATCGCGTCGGCCAGATCCCAGCCGTGA
TACATGCTTTCCCACTCGCGATGACCGCTGAAGCGACCCATCGCCGGGGTCG
GCCGGCCCTGATATTCGCCCTGGAAGGCGACCTTGTGGGTCCAACGGTCCAC
TTCATGGCCGTAGGTGTAGATTTCGCCGTCCACCTCATCGACCACGATATCCT
CGCGGATCAGGCAGGGAACGAGGTTGGACCAGCAGCGATGCGGATAGACGT
AGCCGGTATCGACGAAGGTGATCGGCGGATTGCCGGGTTTCGACAGCTTGGC
ATAGTTTTCCCACCAAATACCGAATTCATCGTACCAGCCGGGATACTTGTGCT
CGAACCACTCGAAGTCACGCTCGGTCATCGCCTCGATGCGCCAGAAGTTGGC
CCACCAGCCGGCGGAGAAGAACTGCGCGACCTTGTGGACGTAGTTCTTCTTG
ACGATCCGGTCGAAGGCTTCGTGGACATCGTCGTAGTGGATCTTGATGCCGT
ATTTCTCCAGCGGCAGCATATAGGTCCGGTAAT

7) COJC7 (3-11-10_COJC_B1)

GGAACGGACGAAACCCATGTCCTTGATCGCGTCGGCCAGATCCCAGCCGTGA
TACATGCTCTCCCACTCGCGATGACCGCTGAAGCGACCCATGGCAGGCGTCG
GACGCCCCTGATACTCGCCCTGGAAGGCGACCTTGTGGGTCCAGCGGTCCAC
TTCATGGCCGTAGGTGTAGATCTCACCATCCACCTCATCGACCACGATGTCCT

CGCGGATCAGGCAGGGAACGAGGTTGGACCAGCAGCGATGCGGATAGACGT
AGCCGGTATCGACGAAGGTGATCGGCGGATTGCCGGGTTTCGACAGCTTGGC
ATAGTTTTCCCACCAGATACCGAATTCATCGTACCAGCCGGGATACTTGTGCT
CGAACCACTCGAAGTCACGCTCGGTCATCGCCTCGATGCGCCAGAAGTTGGC
CCACCAGCCGGCGGAGAAGAACTGGGCGACCTTGTGGACGTAGTTCTTCTTG
ACGATCCGGTCGAAAGCTTCGTGGACATCGTCGTGGTGGATCTTGATGCCGT
ACTTCTCAAGCGGCAGCATATAGGTCCGATAGT

8) COJC8 (3-11-10_COJC_B2)

GGAACGGACGAAACCCATGTCCTTGATCGCGTCGGCCAGATCCCAGCCGTGA
TACATGCTCTCCCACTCGCGATGACCGCTGAAGCGACCCATGGCAGGCGTCG
GACGCCCCTGATACTCGCCCTGGAAGGCGACCTTGTGGGTCCAGCGGTCCAC
TTCATGGCCGTAGGTGTAGATCTCGCCATCCACCTCATCGACCACGATGTCCT
CGCGGATCAGGCAGGGAACGAGGTTGGACCAGCAGCGATGCGGATAGACGT
AGCCGGTATCGACGAAGGTGATCGGCGGATTGCCGGGTTTCGACAGCTTGGC
ATAGTTTTCCCACCAGATACCGAATTCATCGTACCAGCCGGGATACTTGTGCT
CGAACCACTCGAAGTCACGCTCGGTCATCGCCTCGATGCGCCAGAAGTTGGC
CCACCAGCCGGCGGAGAAGAACTGGGCGACCTTGTGGACGTAGTTCTTCTTG
ACGATCCGGTCGAAAGCTTCGTGGACATCGTCGTGGTGGATCTTGATGCCGT
ACTTCTCAAGCGGCAGCATATAGGTCCGATAGT

9) COJC9 (3-11-10_COJC_B3)

GGAACGGACGAAACCCATGTCCTTGATCGCGTCGGCCAGATCCCAGCCGTGA
TACATGCTTTCCCACTCGCGATGACCGCTGAAGCGGCCCATCGCCGGGGTCG
GCCGTCCCTGGTATTCGCCCTGGAAAGCGACCTTGTGGGTCCAGCGGTCGAC
TTCATGGCCGTAGGTGTAGATCTCGCCGTCCACCTCATCGACCACGATGTCCT
CGCGGATCAGGCAGGGAACGAGGTTCGACCAGCAGCGATGTGGATAGACGT
AGCCGGTATCGACGAAGGTGATCGGCGGATTGCCGGGTTTCGACAGCTTGGC
ATAGTTCTCCCACCACACGCCGAACTCATCGTACCAGCCGGGATACTTGTGCT
CGAACCACTCGAAGTCACGCTCGGTCATCGCCTCGATGCGCCAGAAGTTGGC
CCACCAGCCGGCGGAGAAGAACTGGGCGACCTTGTGGACGTAGTTCTTCTTG
ACGATCCGGTCGAAAGCTTCATGGACATCGTCATGGTGGATCTTGATGCCGT
ATTTCCCCAGCGGCAGCATATAAGTCCGGTAAT

10) COJC10 (3-11-10_COJC_C1)

GTTGCGGACGTAGCCAAGATCCTTAAAACACGTCCTCCAGGTCCCCAACCCG
TGGTACAGGGTCTCCCACTCGCGCTTGCCGGTGAGCCGGCCCATGGAGGGCG
TCGGACGGCCGTTGTACTCCTCGCGGAAGGCGACTTTGTCCGTCCAGTGGCA
AGTCTCCGAGCAGTAGGTCCGCCACTCGCCGTCAACGTGGTCGAGGACCGTG
TCCTCGCGGATCAGGCACGGCACCATGCAGGTCCAGCAGCGGTTCGGATACC
AGTAGCCGGTGTCCTCGAACGCTATAGGCTTGTGGCCGTTAGGCTCAGAGAA
ATTCCGGTAGTGCTCCCACCACTTGCCGAACTTGTCGTACCAGCCGGGGTACT
TGTGCTCGAACCACTCGAAGTCTTCCTCGGTCATCGGGTCGATGCGCCAGTAG
TTGGCGAACCAGCCGGTGGCGAAGAACTGGGCCACATAGTGGACGTACCACT

TGTCCCAGACCCGGTTCCACGACTCCTCGATCAGGCCGTGCGGGACCTCCAG
GCCGTACTTCTCGAGCGGGACCAGGTAGCTGCGATAGT

11) COJC11 (3-11-10_COJC_C3)

GGAACGGACGAAACCCATGTCCTTGATCGCGTCGGCCAGATCCCAGCCGTGA
TACATGCTCTCCCACTCGCGATGACCGCTGAAGCGGCCCATCGCCGGCGTCG
GAAGCCCCTGATATTCGCCTTGGAAAGCGACCTTGTGGGTCCAGCGGTCGAC
TTCATGGCCGTAGGTGTAGATCTCGCCGTCCACCTCATCGACCACGATGTCCT
CGCGGATCAGGCAGGGAACGAAGTTGGACCAGCAGCGATGCGGATAGACGT
AGCCGGTATCGACGAAGGTGATCGGCGGATTGCCGGGTTTCGACAGCTTGGC
ATAGTTTTCCCACCAAATACCGAATTCATCGTACCAGCCGGGATACTTGTGCT
CGAACCACTCGAAGTCACGCTCGGTCATCGCCTCGATGCGCCAGAAGTTGGC
CCACCAGCCGGCGGAGAAGAACTGGGCGACTTTGTGGACGTAGTTCTTCTTG
ACGATCCGGTCGAAAGCTTCGTGGACATCGTCGTGGTGGATCTTGATGCCGT
ACTTCTCCAGCGGCAACATGTAGGTCCGGTAAT

12) COJC12 (3-11-10_COJC_D2)

GNAACGGACGAAACCCATGTCCTTGATCGCGTCGGCNANATCCCAGCCGTGA
TACATGCTTTCCCACTCGCGATGACCGNTGAAGCGACCCATCNCCGGGGTCG
GCCGGCCNTGATATTCGCCCTGGAAGGCGACCTTGNGGGTCCAACGGTNCAC
TTCNTGNCCGTAGGTGTAGATTTCGCCGTCCACCTCATCGACCNCGATATCCT
CGNGGATCAGGCAGGGAACGAGGTTGGACCANNAGCGATGCGGATAGACGT
ANCCGGTATCGACNAAGGTGATCGGCGGATTGCCGGGTTTCGACAGCTTGGC
ATAGTTTTCCCACCAAATACNGAATTCATCGCACCAGCCGGGATACTTGTGCT
CGAACCACTNGAAGTCACGNTCGGTCATCGCCTCGATGCGCCAGAAGTTGGC
CCACCAGCCGGCGGANAAGAACTGCGCGACCTNGTGGACGTAGTTCTTCTTG
ACGATCCGGTCGAAGGCTTNNTGGACATNGTCGTGGTGGATCTTGATGCCNT
ATTTCTCCAGCGGCANCATATAGGTCCTGTAAT

13) COJC13 (3-11-10_COJC_F1)

GGAACGGACGAAACCCATGTCCTTGATCGCGTCGGCCAGATCCCAGCCATGA
TACATGCTCTCCCACTCGCGATGACCGCTGAAGCGGCCCATGGCCGGGGTCG
GCCGGCCCTGGTATTCGCCCTGGAAAGCGACCTTGTGGGTCCAGCGATCGAC
TTCATGGCCGTAGGTGTAGATCTCGCCGTCCACCTCATCGACTACGATGTCCT
CGCGGATCAGGCAGGGAACGAGGTTGGACCAGCAGCGATGCGGATAGACGT
AGCCGGTATCGACGAAGGTGATCGGCGGGTTACCGGGCTTCGACAGCTTGGC
GTAGTTCTCCCACCAGATACCGAATTCATCGTACCAGCCGGGATACTTGTGCT
CGAACCACTCGAAGTCACGCTCGGTCATGGCCTCGATGCGCCAGAAGTTGGC
CCACCAGCCGGCGGAGAAGAACTGGGCGACCTTGTGGACGTAGTTTTCTTG
ACGATCCGGTCGAAAGCTTCATGGACATCGTCATGGTGGATCTTGATGCCGT
ACTTCTCCAGCGGCAGCATGTAGGTCCGGTAAT

14) COJC14 (3-11-10_COJC_G1)

GGGCCGGACGAAGCCCATGTCCTTGATCGCGTCGGCCAGATCCCAGCCGTGA
TACATGCTTTCCCATTCGCGATGACCACTGAAGCGGCCCATCGCTGGGGTTGG
GCGGCCCTGATATTCACCTTGGAAGGCAACCTTGTGAGTCCACCTGTCCACTT
CGTGCCCATAGGTGTAGATCTCGCCGTCGACCTCATCGACCACGATGTCCTCA
CGGATCAGGCAGGGCACCAGGTTCGACCAGCAGCGATGCGGATAGACATAG
CCGGTATCGACGAAGGTGATCAGCGGGTTGCCGGGCTTGCTGAGCTTGGCAT
AGTTTTCCCACCACGCGCCGAATTCATCGTACCAACCCGGATACTTGTGCTCG
AACCACTCGAAGTCCCGCTCCGTCATGGCCTCGATGCGCCAGAAGTTGGCCC
ACCAGCCGGCCGAGAAGAACTGGGCGACCTTATGGACATAGTTCTTCTTGAC
GATCCGGTCGAACGCTTCGTGGACATCGTCATGGTGGATCTTGATGCCGTACT
TCTCCAGCGGCAGCATGTAGGTCCGATAAT

15) COJC15 (3-11-10_COJC_G2)

GGAACGGACGAAACCCATGTCCTTGATCGCGTCGGCCAGATCCCAGCCGTGA
TACATGCTCTCCCACTCGCGATGACCGCTGAAGCGGCCCATCGCCGGGGTCG
GCCGGCCCTGGTATTCGCCCTGGAAGGCGACCTTATGGGTCCAGCGGTCCAC
TTCATGGCCGTAGGTGTAGATCTCGCCGTCCACCTCATCGACCACGATGTCCT
CGCGGATCAGGCAGGGAACGAGATTCGACCAGCAGCGATGCGGATAAACAT
AGCCGGTATCGACGAAGGTGATCGGCGGGTTGCCGGGCTTCGACAGCTTGGC
ATAGTTTTCCCACCAGATACCGAACTCATCGTACCAGCCGGGATACTTGTGCT
CGAACCACTCGAAATCACGCTCGGTCATAGCCTCGATGCGCCAGAAGTTGGC
CCACCAGCCGGCGGAGAAGAACTGGGCGACTTTGTGAACGTAATTCTTCTTG
ACGATCCGGTCGAAAGCTTCGTGGACATCGTCGTGGTGGATCTTGATGCCGT
ATTTCTCCAGCGGCAGCATGTAGGTCCGGTAAT

16) COJC16 (3-11-10_COJC_H2)

GGAACGGACNAAANNCATGTCCNTGATCGCGTCGGCCANACCCCANCCGTG
ATACATGCTTTCCCACTCNCGATGACCGCTGAAGCGACCCATCGCCGGGGTC
GGCCGGCCCTGATATTCGCCCTGGAAGGCGACCTTGTGGGTCCAACGGTCCA
CTTCATGGCCGTAGGTGTAGATTTCGCCGTCCACCTCATCGACCANGATATCC
TNGNGGATNAGGCAGGGAACGAGGNTGGACCANCAGCNATGCGGATAGACG
TAGCCGGTATCGACGAAGGNGATCGGCGGATTGCCGGGTTTNGATAGCNTGG
CATAGTTTTCCCACCAAATACCGAATTCATNGTACCAGCCGGGATACTTGTGC
TCNAACCACTCNAAGTCACGCTCGGTCATCGCCTCGATGCGCCAGAAGTTGG
CCCACCAGCCGGCGGAGAANAANTGCGCGACCTTGTGGACGTAGTTCGTCTT
GACNATCCGGTNGAAGGCTTCGNGGACATCNTCNTGGTGGATCTTNATGCCN
NATTTCTTCNGCGGCNGCATATAGGTCCGGTAAT

17) COJC17 (3-11-10_COJC_H4)

CGGGCGGACGAAGCCCATGTCCTTGTTCGCGTCGGCCAGATCCCAGCCGTGA
TACATGCTTTCCCACTCACGATGACCGCTGAAGCGGCCCATCGCCGGGGTCG
GGCGGCCTTGATATTCGCCCTGGAAAGCGACCTTGTGGGTCCACCGGTCGAC
TTCGTGGCCATAGGTGTAGATCTCGCCGTCGACTTCATCGACCACGATATCCT

CACGGATTAGGCAGGGCACCAGATTCGACCAGCAGCGATGCGGATAGACAT
AGCCGGTATCGACGAAGGTGATCGGCGGGTTGCCGGGCTTGCTGAGCTTAGC
ATAGTTCTCCCACCACGCACCGAATTCATCGTACCAACCGGGATACTTGTGCT
CGAACCATTCGAAGTCCCGCTCCGTCATGGCCTCGATGCGCCAGAAGTTGGC
CCACCAGCCAGCCGAGAAGAACTGCGCTACCTTATGGACATAGTTCTTCTTG
ACGATCCGGTCGAACGCTTCGTGGACATCGTCATGGTGGATCTTGATGCCGTA
CTTCTCCAGCGGCAGCATGTAGGTCCGGTAAT

18) COJC18 (5-25-10_COJC_6-1-10-1)

GGAACGGACGAAACCCATGTCCTTGATCGCGTCGGCCAGATCCCAGCCGTGA
TACATGCTTTCCCACTCGCGATGACCGCTGAAGCGACCCATCGCCGGGGTCG
GCCGGCCCTGATATTCGCCCTGGAAGGCGACCTTGTGGGTCCAACGGTCCAC
TTCATGGCCGTAGGTGTAGATTTCGCCGTCCACCTCATCGACCACGATATCCT
CGCGGATCAGGCAGGGAACGAGGTTGGACCAGCAGCGATGCGGATAGACGT
AGCCGGTATCGACGAAGGTGATCGGCGGATTGCCGGGTTTCGACAGCTTGGC
ATAGTTTTCCCACCAAATACCGAATTCATCGTACCAGCCGGGATACTTGTGCT
CGAACCACTCGAAGTCACGCTCGGTCATCGCCTCGATGCGCCAGAAGTTGGC
CCACCAGCCGGCGGAGAAGAACTGCGCGACCTTGTGGACGTAGTTCTTCTTG
ACGATCCGGTCGAAGGCTTCGTGGACATCGTCGTGGTGGATCTTGATGCCGT
ATTTCTCCAGCGGCAGCATATAGGTCCGGTAAT

19) COJC19 (5-25-10_COJC_6-1-10-2)

GGAACGGACGAAACCCATGTCCTTGATCGCGTCGGCCAGATCCCAGCCGTGA
TACATGCTTTCCCACTCGCGATGACCGCTGAAGCGACCCATCGCCGGGGTCG
GCCGGCCCTGATATTCGCCCTGGAAGGCGACCTTGTGGGTCCAACGGTCCAC
TTCATGGCCGTAGGTGTAGATTTCGCCGTCCACCTCATCGACCACGATATCCT
CGCGGATCAGGCAGGGAACGAGGTTGGACCAGCAGCGATGCGGATAGACGT
AGCCGGTATCGACGAAGGTGATCGGCGGATTGCCGGGTTTCGACAGCTTGGC
ATAGTTTTCCCACCAAATACCGAATTCATCGTACCAGCCGGGATACTTGTGCT
CGAACCACTCGAAGTCACGCTCGGTCATCGCCTCGATGCGCCAGAAGTTGGC
CCACCAGCCGGCGGAGAAGAACTGCGCGACCTTGTGGACGTAGTTCTTCTTG
ACGATCCGGTCGAAGGCTTCGTGGACATCGTCGTGGTGGATCTTGATGCCGT
ATTTCTCCAGCGGCAGCATATAGGTCCGGTAAT

20) COJC20 (5-25-10_COJC_6-1-10-3)

TGAACGGACGAAACCCATGTCCTTGATCGCGTCGGCCAGATCCCACCCGTGA
TACATGCTCTCCCACTCGCGATGACCGCTGAAGCGGCCCATGGCCGGGGTCG
GTCGACCCTGATACTCGCCCTGGAAGGCGACCTTGTGGGTCCAACGGTCGAC
TTCATGGCCGTAGGTGTAGATCTCGCCGTCCACCTCATCGACCACGATGTCCT
CGCGGATCAGGCAGGGAACGAGGTTGGACCAGCAGCGATGCGGATAAACGT
AGCCGGTATCGACGAAGGTGATCGGCGGGTTGCCGGGCTTCGACAGCTTGGC
ATAGTTCTCCCACCAGATACCGAATTCATCGTACCAGCCGGGATACTTGTGCT
CGAACCACTCGAAGTCACGCTCGGTCATCGCCTCGATGCGCCAGAAGTTGGC
CCACCAGCCGGCGGAGAAGAACTGAGCGACCTTGTGGACGTAGTTCTTCTTG

ACGATCCGGTCGAACGCTTCGTGGACATCGTCGTGGTGGATCTTGATGCCGTA
CTTCTCCAGCGGCAGCATGTAGGTCCGGTAAC

21) COJC21 (5-25-10_COJC_6-1-10-4)

GGAACGGACGAAACCCATGTCCTTGATCGCGTCGGCCAGATCCCAGCCGTGA
TACATGCTTTCCCACTCGCGATGACCGCTGAAGCGACCCATCGCCGGGGTCG
GCCGGCCCTGATATTCGCCCTGGAAGGCGACCTTGTGGGTCCAACGGTCCAC
TTCATGGCCGTAGGTGTAGATTTCGCCGTCCACCTCATCGACCACGATATCCT
CGCGGATCAGGCAGGGAACGAGGTTGGACCAGCAGCGATGCGGATAGACGT
AGCCGGTATCGACGAAGGTGATCGGCGGATTGCCGGGTTTCGACAGCTTGGC
ATAGTTTTCCCACCAAATACCGAATTCATCGTACCAGCCGGGATACTTGTGCT
CGAACCACTCGAAGTCACGCTCGGTCATCGCCTCGATGCGCCAGAAGTTGGC
CCACCAGCCGGCGGAGAAGAACTGCGCGACCTTGTGGACGTAGTTCTTCTTG
ACGATCCGGTCGAAGGCTTCGTGGACATCGTCGTGGTGGATCTTGATGCCGT
ATTTCTCCAGCGGCAGCATATAGGTCCGGTAAT

22) COJC22 (5-25-10_COJC_6-1-10-6)

GGAACGGACGAAACCCATGTCCTTGATCGCGTCGGCCAGATCCCAGCCGTGA
TACATGCTTTCCCACTCGCGATGACCGCTGAAGCGGCCCATCGCCGGGGTCG
GCCGTCCCTGATATTCGCCCTGGAAGGCGACTTTATGGGTCCAGCGGTCCACT
TCATGGCCGTAGGTCTAGATCTCACCGTCCACCTCATCGACCACGATGTCCTC
GCGGATCAGGCAGGGAACGAGGTTGGACCAGCAGCGATGCGGATAAACGTA
GCCGGTATCGACGAAGGTGATCGGCGGGTTGCCGGGCTTCGACAGCTTGGCA
TAGTTCTCCCACCAGATACCGAATTCATCGTACCAGCCGGGATACTTGTGCTC
GAACCACTCGAAGTCACGCTCGGTCATCGCCTCGATGCGCCAGAAGTTGGCC
CACCAGCCGGCGGAGAAGAACTGGGCGACCTTGTGGACGTAGTTCTTCTTGA
CGATCCGGTCGAAGGCTTCGTGGACATCGTCGTGGTGGATCTTGATGCCGTAT
TTCTCCAGCGGCAGCATATAGGTCCGGTAAT

23) COJC23 (5-25-10_COJC_6-1-10-9)

GTCACGCACGTACCCCATGTCGGAGACGACGTCGGCCCAGTTCCAGCCGTGG
TACAGCGTCTCCCACTCGCGCTTGCCGATGAGCTGACCCATGTTCGGCGTCTC
GCGGCCCTGGTAGGTGGGCCGGAAGGCGACCGTGTCGGTCCAGCGGCATGCC
TCGTGGCAGTACGTCCGCGTCTGCCCGTCGACCTCGGCCATGACCATGTCCTC
ACGGACGAGGCACGGCACCATGCAGGTCCAGCACCGGGCCGGGTAGACGTA
GTTGACGTCCTCCAGCGCGATGGGGTTGTGCCCGTTGGCGACCGACAACCGG
GAGTAGTTCTCCCACCAGGCGCCGTACTTGTCGTACCAGCCGGGGTACTTGTA
CTCGAACCACTCGAAGTCCTCGTCGGTCATCGGGTCGATCCGCCAGTAGTTG
GCCAGCCAGCCCGTGGCAAAGAACTGCGCCACCTCGTGCACGTAGCCCTTGT
TCCAGATCTGGTTCCACGACTCCTCGATGAGGTCGTGCGGGATCACCAGGCC
GTACTTCTCCAGCGGGACCAGGTAGCTGCGGTAGT

24) COJC24 (5-25-10_COJC_6-1-10-10)

GGAACGGACGAAACCCATGTCCTTGATCGCGTCGGCCAGATCCCAGCCGTGA
TACATGCTCTCCCACTCGCGATGACCGCTGAAGCGACCCATGGCAGGCGTCG
GACGCCCCTGATACTCGCCCTGGAAGGCGACCTTGTGGGTCCAGCGGTCCAC
TTCATGGCCGTAGGTGTAGATCTCGCCATCCACCTCATCGACCACGATGTCCT
CGCGGATCAGGCAGGGAACGAGGTTGGACCAGCAGCGATGCGGATAGACGT
AGCCGGTATCGACGAAGGTGATCGGCGGATTGCCGGGTTTCGACAGCTTGGC
ATAGTTTTCCCACCAGATACCGAATTCATCGTACCAGCCGGGATACTTGTGCT
CGAACCACTCGAAGTCACGCTCGGTCATCGCCTCGATGCGCCAGAAGTTGGC
CCACCAGCCGGCGGAGAAGAACTGGGCGACCTTGTGGACGTAGTTCTTCTTG
ACGATCCGGTCGAAAGCTTCGTGGACATCGTCGTGGTGGATCTTGATGCCGT
ACTTCTCAAGCGGCAGCATATAGGTCCGATAGT

25) COJC25 (5-25-10_COJC_6-1-10-11)

GGAACGGACGAAACCCATGTCCTTGATCGCGTCGGCCAGATCCCAGCCGTGA
TACATGCTCTCCCACTCGCGATGACCGCTGAAGCGGCCCATGGCGGGTGTCG
GGCGCCCCTGATATTCGCCCTGGAAAGCGACCTTGTGGGTCCAGCGATCGAC
TTCATGGCCGTAGGTGTAGATTTCGCCATCCACCTCATCGACCACGATGTCTT
CGCGGATCAGGCAGGGAACGAGGTTCGACCAGCAGCGGTGCGGATAGACGT
AGCCGGTATCGACAAAGGTAATCGGCGGGTTGCCGGGCTTCGACAGCTTGGC
ATAGTTTTCCCACCAGGTGCCGAATTCATCGTACCAGCCGGGATACTTGTGCT
CGAACCACTCGAAGTCACGCTCGGTCATCGCCTCGATGCGCCAGAAGTTGGC
CCANCAGCCGGCCGAGAAGAACTGTGCGACCTTGTGGACNTANTTCTTTCTT
GACGATCCGATCGAAGGCTTCGTGGACGTCGTCGTGATGGATCCTGATGCCG
TACTTCTCCAGCGGCAGCATATAGGTCACGGNAAT

26) COJC26 (5-25-10_COJC_6-1-10-14)

GGAACGGACGAAACCCATGTCCTTGATCGCGTCGGCCAGATCCCAGCCGTGA
TACATGCTNTCCCACTCGCGATGACCGCTGAAGCGNCCCATNGCNGGNGTCG
GNCGNCCCTGATATTCGCCCTGGAANGCGACCTTGTGGGTCCANCGNTCNAC
TTCATGGCCGTAGGTGTAGATTTCGCCNTCCACCTCATCGACCACGATNTCNT
CGCGGATCAGGCAGGGAACGAGGTTNGACCAGCAGCGNTGCGGATAGACGT
AGCCGGTATCGACNAAGGTNATCGGCGGNTTGCCGGGNTTCGACAGCTTGGC
ATAGTTTTCCCACCANNTNCCGAATTCATCGTACCAGCCGGGATACTTGTGCT
CGAACCACTCGAAGTCACGCTCGGTCATCGCCTCGATGCGCCAGAAGTTGGC
CCACCAGCCGGCNGAGAAGAACTGNGCGACCTTGTGGACGTAGTTCTTCTTG
ACGATCCGNTCGAAGGCTTCGTGGACNTCGTCGTGNTGGATCNTGATGCCGT
ANTTCTCCAGCGGCAGCATATAGGTCCGGTAAT

27) COJC27 (5-25-10_COJC_6-1-10-15)

GGAACGGACGAAACCCATGTCCTTGATCGCGTCGGCCAGATCCCAGCCGTGG
TACATGCTCTCCCACTCGCGATGACCGCTGAAGCGGCCCATCGCCGGGGTCG
GCCGGCCCTGGTATTCGCCCTGGAAAGCGACCTTGTGGGTCCAGCGATCGAC
TTCATGGCCGTAGGTGTAGATCTCGCCGTCCACCTCATCGACCACGATGTCCT

CGCGGATCAGGCAGGGAACGAGGTTGGACCAGCAGCGATGCGGATAGACGT
AGCCAGTATCGACGAAGGTGATCGGCGGATTGCCGGGCTTCGACAGCTTGGC
GTAGTTCTCCCACCAGATACCGAATTCATCGTACCAGCCGGGATACTTGTGCT
CGAACCACTCAAAGTCACGTTCGGTCATCGCCTCGATGCGCCAGAAGTTGGC
CCACCAGCCGGCGGAGAAGAACTGGGCGACCTTGTGGACATAGTTCTTCTTG
ACGATCCGGTCGAAGGCTTCGTGGACATCGTCGTGGTGGATCTTGATGCCGT
ACTTCTCCAGCGGCAGCATGTAGGTCCGGTAGT

28) COJC28 (5-25-10_COJC_A1)

GGACCGGACGAAGTTCANGTCCTTGATGGCGTCGGCGAGATCCCAGCCGTGG
TACAGCGTCTCCCACTCGCGCTTGCCGGAGAACCGGCCCATCGCGGGCGTCG
GACGGCCCTGATACTCGTCCGCGAAGGCCTCGACGGCAGTCCAGCGGTCGAG
CTCGTGGGCGAAGGTGTGGAGCTTGCCGTCNATCTCGTCCACCACCATGTCCT
CGCGGATCAGGCACGGGACCAGGCACGACCAGCAGCGGTGCGGATAGACGT
AGCCGGTGATGTCCTGCGCGAAGGTGACGACCTTGCTGCCCGGCTTGGAGAG
CTTGTCGTACCACTTCCNGAAATCGCCAAACTCGGCATACCAGCCCGGATAC
TTGTGCTCGAACCACTCGAAGTCGGCATCGCGCTGGGCCTCGATGCGCCAGA
AATTGACCGGCCAGCCGACGGCGAAGAACTGCGCCACCTTGTGCACGTAGAA
CTTCTCGGTGATGCGCTTCCAGGCCGNNTGGACGTCGTCGTGATGGATCTTGA
TGCCGTATTTCTCCANCGGCAGCATGTAGGTGCGGTAGT

29) COJC29 (5-25-10_COJC_A2)

GGAACGGACGAAACCCATCCACCGGTGCCAGTCCTTGATCGCGTCGGCCAGA
TCCCAGCCGTGATACATGCTTTCCCACTCGCGATGACCGCTGAAGCGACCCAT
CGCCGGGGTCGGCCGGCCCTGATATTCGCCCTGGAAGGCGACCTTGTGGGTC
CAACGGTCCACTTCATGGCCGTAGGTGTAGATTTCGCCGTCCACCTCATCGAC
CACGATATCCTCGCGGATCAGGCAGGGAACGAGGTTGGACCAGCAGCGATG
CGGATAGACGTAGCCGGTATCGACGAAGGTGATCGGCGGATTGCCGGGTTTC
GACAGCTTGGCATAGTTTTCCCACCAAATACCGAATTCATCGTACCAGCCGG
GATACTTGTGCTCGAACCACTCGAAGTCACGCTCGGTCATCGCCTCGATGCGC
CAGAAGTTGGCCCACCAGCCGGCGGAGAAGAACTGCGCGACCTTGTGGACGT
AGTTCTTCTTGACGATCCGGTCGAAGGCTTCGTGGACATCGTCGTGGTGGATC
TTGATGCCGTATTTCTCCAGCGGCAGCATATAGGTCCGGTAAT

30) COJC30 (5-25-10_COJC_A4)

GGAACGGACGAAACCCATGTCCTTGATCGCGTCGGCCAGATCCCAGCCGTGA
TACATGCTTTCCCACTCGCGATGACCGCTGAAGCGACCCATCGCCGGGGTCG
GCCGGCCCTGATATTCGCCCTGGAAGGCGACCTTGTGGGTCCAACGGTCCAC
TTCATGGCCGTAGGTGTAGATTTCGCCGTCCACCTCATCGACCACGATATCCT
CGCGGATCAGGCAGGGAACGAGGTTGGACCAGCAGCGATGCGGATAGACGT
AGCCGGTATCGACGAAGGTGATCGGCGGATTGCCGGGTTTCGACAGCTTGGC
ATAGTTTTCCCACCAAATACCGAATTCATCGTACCAGCCGGGATACTTGTGCT
CGAACCACTCGAAGTCACGCTCGGTCATCGCCTCGATGCGCCAGAAGTTGGC
CCACCAGCCGGCGGAGAAGAACTGCGCGACCTTGTGGACGTAGTTCTTCTTG

ACGATCCGGTCGAAGGCTTCGTGGACATCGTCGTGGTGGATCTTGATGCCGT
ATTTCTCCAGCGGCAGCATATAGGTCCGGTAAT

31) COJC31 (5-25-10_COJC_A5)

GGAACGGACGAAACCCATGTCCTTGATCGCGTCGGCCAGATCCCAGCCGTGA
TACATGCTCTCCCACTCGCGATGGCCGCTGAAGCGACCCATCGCCGGGGTCG
GCCGTCCCTGGTATTCGCCCTGGAAGGCGACCTTGTGGGTCCAGCGGTCCACT
TCGTGGCCGTAGGTGTAGATTTCGCCGTCCACCTCATCGACCACGATGTCCTC
GCGGATCAGGCAGGGAACGAGGTTCGACCAGCAGCGGTGCGGATAGACGTA
GCCGGTATCGACGAAGGTGATCGGCGGATTGCCGGGCTTCGACAGCTTGGCG
TAGTTCTCCCACCAGATACCGAATTCATCGTACCAGCCGGGATACTTGTGCTC
GAACCACTCGAAGTCACGCTCGGTCATCGCCTCGATGCGCCAGAAGTTGGCC
CACCAGCCGGCGGAAAAGAACTGGGCGACTTTGTGGACGTAGTTCTTCTTGA
CGATCCGGTCGAACGCTTCATGGACATCGTCGTGGTGGATCTTGATGCCGTAC
TTCTCCAGCGGCAGCATATAGGTCCGGTAAT

32) COJC32 (5-25-10_COJC_A6)

GGAACGGACGAAACCCATGTCCTTGATCGCGTCGGCCAGATCCCAGCCGTGA
TACATGCTTTCCCACTCGCGATGACCGCTGAAGCGACCCATCGCCGGGGTCG
GCCGGCCCTGATATTCGCCCTGGAAGGCGACCTTGTGGGTCCAACGGTCCAC
TTCATGGCCGTAGGTGTAGATTTCGCCGTCCACCTCATCGACCACGATATCCT
CGCGGATCAGGCAGGGAACGAGGTTGGACCAGCAGCGATGCGGATAGACGT
AGCCGGTATCGACGAAGGTGATCGGCGGATTGCCGGGTTTCGACAGCTTGGC
ATAGTTTTCCCACCAAATACCGAATTCATCGTACCAGCCGGGATACTTGTGCT
CGAACCACTCGAAGTCACGCTCGGTCATCGCCTCGATGCGCCAGAAGTTGGC
CCACCAGCCGGCGGAGAAGAACTGCGCGACCTTGTGGACGTAGTTCTTCTTG
ACGATCCGGTCGAAGGCTTCGTGGACATCGTCGTGGTGGATCTTGATGCCGT
ATTTCTCCAGCGGCAGCATATAGGTCCGGTAATAG

33) COJC33 (5-25-10_COJC_A7)

GGAACGGACGAAACCCATGTCCTTGATCGCGTCGGCCAGATCCCAGCCATGA
TACATGCTCTCCCACTCGCGATGACCGCTGAAGCGGCCCATGGCCGGGGTCG
GCCGTCCCTGATATTCGCCTTGGAAAGCGACCTTGTGGGTCCAGCGATCGACT
TCATGGCCGTAGGTGTAGATCTCGCCATCGACCTCATCGACCACGATATCCTC
GCGGATCAGACAGGGAACGAGGTTCGACCAGCAGCGGTGCGGATAGACGTA
GCCGGTATCGACGAAGGTGATCGGTGGATTGCCGGGCTTTGACAGCTTGGCA
TAATTTTCCCACCACACGCCGAATTCATCGTACCAGCCCGGATACTTGTGCTC
GAACCACTCGAAGTCACGCTCGGTCATCGCCTCGATGCGCCAGAAGTTGGCC
CACCAGCCGGCCGAAAAGAACTGGGCGACCTTGTGGACGTAGTTCTTCTTGA
CGATCCGGTCGAACGCTTCGTGGACATCGTCATGGTGGATCCTGATGGCGTA
CTTCTCCAGCGGCAGCATATAGGTTCGATAGT

34) COJC34 (5-25-10_COJC_A10)

GGAACGGACGAAACCCATGTCTTTGATCGCGTCGGCCAGATCCCAGCCGTGA
TACATGCTCTCCCACTCGCGATGACCGCTGAAGCGGCCCATCGCCGGGGTCG
ACCGGCCCCGGTATTCGCCCTGGAAGGCGACCTTATGGGTCCAGCGGTCCAC
TTCATGGCCGTAGGTGTAGATCTCGCCGTCCACCTCATCGACCACGATGTCCT
CGCGGATCAGGCAGGGAACGAGGTTGGACCAGCAGCGATGCGGATAAACAT
AGCCGGTATCGACGAAGGTGATCGGCGGGTTGCCGGGCTTCGACAGCTTGGC
GTAGTTCTCCCACCAGATACCGAATTCATCGTACCAGCCCGGATACTTGTGCT
CGAACCACTCGAAGTCACGCTCGGTCATCGCCTCGATGCGCCAGAAGTTGGC
CCACCAGCCGGCGGAGAAGAACTGGGCGACCTTGTGGACGTAGTTCTTCTTG
ACGATCCGGTCGAAAGCTTCATGGACATCGTCATGGTGGATCTTGATGCCGT
ATTCCTCCAGCGGCAGCATATAGGTCCGGTAAT

35) COJC35 (5-25-10_COJC_N1)

GGAACGGACGAAACCCATGTCCTTGATCGCGTCGGCCAGATCCCAGCCGTGA
TACATGCTCTCCCACTCGCGATGACCGCTGAAGCGACCCATGGCAGGCGTCG
GACGCCCCTGATACTCGCCCTGGAAGGCGACCTTGTGGGTCCAGCGGTCCAC
TTCATGGCCGTAGGTGTAGATCTCGCCATCCACCTCATCGACCACGATGTCCT
CGCGGATCAGGCAGGGAACGAGGTTGGACCAGCAGCGATGCGGATAGACGT
AACCGGTATCGACGAAGGTGATCGGCGGATTGCCGGGTTTCGACAGCTTGGC
ATAGTTTTCCCACCAGATACCGAATTCATCGTACCAGCCGGGATACTTGTGCT
CGAACCACTCGAAGTCACGCTCGGTCATCGCCTCGATGCGCCAGAAGTTGGC
CCACCAGCCGGCGGAGAAGAACTGGGCGACCTTGTGGACGTAGTTCTTCTTG
ACGATCCGGTCGAAAGCTTCGTGGACATCGTCGTGGTGGATCTTGATGCCGT
ACTTCTCAAGCGGCAGCATATTAGGTCCGATAGT

36) COJC36 (5-25-10_COJC_N2)

TGAACGGACGAAACCCATGTCCTTGATCGCGTCGGCCAGATCCCAGCCGTGA
TACATGCTCTCCCACTCGCGATGACCGCTGAATCGGCCCATGGCCGGGGTCG
GACGCCCTTCGTATTCACCCTGGAAGGCGACCTTGTGGGTCCAGCGGTCCACT
TCATGGCCGTAGGTGTAGATCTCGCCGTCCACCTCATCGACCACGATATCCTC
GCGGATCAGGCAGGGAACGAGGTTCGACCAGCAGCGATGTGGATAGACGTA
GCCGGTATCCACGAAGGTGATCGGCGGGTTGCCGGGCTTCGACAGCTTGGCA
TAGTTTTCCCACCATGCGCCGAATTCATCGTACCAGCCGGGATACTTGTGCTC
GAACCACTCGAAGTCACGCTCGGTCATCGCCTCGATGCGCCAGAAGTTGGCC
CACCAGCCGGCGGAGAAGAACTGGGCGACCTCGTGGACGTAGTTCTTCTTGA
CGATCCGGTCGAATGCTTCGTGGACATCGTCGTGGTGGATCTTGATGCCGTAT
TTCTCCAGCGGCAGCATATAGGTCCGGTAAT

37) COJC37 (5-25-10_COJC_N4)

GGAACGGACGAAACCCATGTCCTTAATCGCGTCGGCCAGATCCCAGCCGTGA
TACATGCTCTCCCACTCGCGATGACCGCTGAAGCGGCCCATCGCCGGGGTCG
GCCGTCCCTGATACTCACCCTGGAAAGCGACCTTGTGGGTCCAGCGATCGAC
TTCATGGCCATACGTGTAGATTTCGCCATCCACCTCATCGACCACGATGTCCT

CGCGGATCAGGCAAGGAACGAGGTTCGACCAGCAGCGGTGCGGATAGACGT
AGCCGGTATCGACGAAGGTGATCGGCGGGTTGCCGGGCTTCGACAGCTTGGC
ATAGTTTTCCCACCAGATACCGAATTCATCGTACCAGCCGGGATACTTGTGCT
CGAACCACTCGAAGTCACGCTCGGTCATCGCCTCGATGCGCCAGAAGTTGGC
CCACCAGCCGGCGGAGAAGAACTGGGCGACCTTGTGGACGTAGTTCTTCTTG
ACGATCCGGTCGAACGCTTCGTGGACATCGTCGTGGTGGATCTTGATGCCGTA
TTTCTCCAGCGGCAGCATATAGGTCCGGTAAT

38) COJC38 (5-25-10_COJC_N5)

CGGGCGGACGAAGCCCATGTCCTTGATCGCGTCGGCCAGATCCCAGCCGTGA
TACATGCTTTCCCACTCACGATGACCGCTGAAGCGGCCCATCGCCGGGGTCG
GGCGGCCTTGATATTCGCCCTGGAAAGCGACCTTGTGGGTCCACCGGTCGAC
TTCGTGGCCATAGGTGTAGATCTCGCCGTCGACTTCATCGACCACGATATCCT
CACGGATTAGGCAGGGCACCAGATTCGACCAGCAGCGATGCGGATAGACAT
AGCCGGTATCGACGAAGGTGATCGGCGGGTTGCCGGGCTTGCTGAGCTTAGC
ATAGTTCTCCCACCACGCACCGAATTCATCGTACCAACCGGGATACTTGTGCT
CGAACCATTCGAAGTCCCGCTCCGTCATGGCCTCGATGCGCCAGAAGTTGGC
CCACCAGCCAGCCGAGAAGAACTGCGCTACCTTATGGACATAGTTCTTCTTG
ACGATCCGGTCGAACGCTTCGTGGACATCGTCATGGTGGATCTTGATGCCGTA
CTTCTCCAGCGGCAGCATGTAGGTCCGGTAAT

39) COJC39 (5-25-10_COJC_N6)

GGAACGGACGAAACCCATGTCCTTGATCGCGTCGGCCAGATCCCAGCCGTGA
TACATGCTTTCCCACTCGCGATGACCGCTGAAGCGACCCATCGCCGGGGTCG
GCCGGCCCTGATATTCGCCCTGGAAGGCGACCTTGTGGGTCCAACGGTCCAC
TTCATGGCCGTAGGTGTAGATTTCGCCGTCCACCTCATCGACCACGATATCCT
CGCGGATCAGGCAGGGAACGAGGTTGGACCAGCAGCAATGCGGATAGACGT
AGCCGGTATCGACGAAGGTGATCGGCGGATTGCCGGGTTTCGACAGCTTGGC
ATAGTTTTCCCACCAAATACCGAATTCATCGTACCAGCCGGGATACTTGTGCT
CGAACCACTCGAAGTCACGCTCGGTCATCGCCTCGATGCGCCAGAAGTTGGC
CCACCAGCCGGCGGAGAAGAACTGCGCGACCTTGTGGACGTAGTTCTTCTTG
ACGATCCGGTCGAAGGCTTCGTGGACATCGTCGTGGTGGATCTTGATGCCGT
ATTTCTCCAGCGGCAGCATATAGGTCCGGTAAT

40) COJC40 (5-25-10_COJC_N7)

GGAACGGACGAAACCCATGTCCTTGATCGCGTCGGCCAGATCCCACCCGTGA
TACATGCTCTCCCACTCGCGATGACCGCTGAAGCGGCCCATCGCCGGGGTCG
ACCGGCCCTGGTATTCGCCCTGGAAGGCGACCTTATGGGTCCAGCGGTCGAC
TTCATGGCCGTAGGTGTAGATCTCGCCATCCACCTCATCGACCACGATGTCCT
CGCGGATCAGGCAGGGAACGAGGTTCGACCAGCAGCGATGCGGATAAACGT
AGCCGGTATCGACGAAGGTGATTGGCGGATTGCCGGGCTTCGACAGCTTGGC
ATAGTTCTCCCACCAGATACCGAACTCATCGTACCAGCCCGGATACTTGTGCT
CGAACCACTCGAAGTCACGCTCGGTCATGGCCTCGATGCGCCAGAAGTTGGC
CCACCAGCCGGCGGAAAAGAACTGGGCGACCTTGTGGACGTAGTTCTTCTTG

ACGATCCGGTCGAACGCTTCGTGGACGTCGTCGTGGTGGATCTTGATGCCGTA
CTTCTCAAGCGGCAGCATATAGGTCCGATAGT

41) COJC41 (5-25-10_COJC_N8)

TGAACGGACGAAACCCATGTCCTTGATCGCGTCGGCCAGATCCCACCCGTGA
TACATGCTCTCCCACTCGCGGTGACCGCTGAAGCGGCCCATGGCCGGGGTCG
GTCGACCCTGATACTCGCCCTGGAAGGCGACCTTGTGGGTCCAACGGTCGAC
TTCATGGCCGTAGGTGTAGATCTCGCCGTCCACCTCATCGACCACGATGTCCT
CGCGGATCAGGCAGGGAACGAGGTTGGACCAGCAGCGATGCGGATAAACGT
AGCCGGTATCGACGAAGGTGATCGGCGGGTTGCCGGGCTTCGACAGCTTGGC
ATAGTTCTCCCACCAGATACCGAATTCATCGTACCAGCCGGGATACTTGTGCT
CGAACCACTCGAAGTCACGCTCGGTCATCGCCTCGATGCGCCAGAAGTTGGC
CCACCAGCCGGCGGAGAAGAACTGGGCGACCTTGTTGGACGTAGTTCTTCTT
GACGATCCGGTCGAACGCTTCGTGGACATCGTCGTGGTGGATCTTGATGCCGT
ACTTCTCCAGCGGCAGCATGTTAGGTCCGGTAAT

42) *M. petroleiphilum* (PCR) 7

CGAGCGGACGAAGCCGAGGTCCTTGATCGCGTCGGCCAGGTCCCAGCCGTGG
TAGCACTCTTCCCATTCGCGGCGACCGCTGAAGCGGCCCATCGCCGGCGTCG
GGCGGCCCTCGTACTCGGCGGCGAAGGCCGTCTTGTGGGTCCAGCGGCACAG
CTCGGAGCAGTAGGTGTAGAGCTTGCCGTCGACCTCGTCGACCACCAGCTCC
TCACGGATCACCGCCGGCACCATGCAGCTCCAGCAGCGGTGCGGGTAGACGT
AGCCGTTCTCCTGGTCGAACAGCATGTTGGTCTCGCCCGGCACGCTCTTGCGC
TCGTACCACTTCCAGAAATCGCCGAACTCGGCGTACCAGCCCGGGTACTTGT
GCTCGAACCACTCGAAGTCCTTCTCGGTCTGGGCTTCGATGCGCCAGAAGTTG
GCGGGCCAGCCGACCGCGAAGAACTGCGCGACCTTGTGCACGTAGTTCTTCT
TGACCAGGCGGTCCCACGCGGCGGAGACGTCGTCGTGGTGGATCTTGACGCC
GTACTTCTCGAGCGGGAGCAGGTAGGTGCGGTAGT

43) *Ps. butanovora* (PCR) E11

GCTGCGCACGCCGCCTGCCGCCTTGACGAGATCGGAGAGTTCCCAGCCTTCG
AACTGCTCGAAGAAGTTCTGGAACTCGTAGCGCTCGGGCTCGAGCAGCCACT
GGCGCTCGCCATACGGCTCGGCGAAGGCGTAGCGCTTGCCGTTGTATTCAGT
GAACCGCGGCTTGATAGAGCTTTTGGCGAGGCTGGGGCAGAACGGCAAACCC
GAGGCATGGTCGACGAAAATGCCGTGACCACGCTCCATCAGCCAGAGCACAC
CACAGAAGCCGCTCTTCGGATCCTCGAAGCCGGCAGCGCGCCACTCCTCGTA
GATCTTGCCGTAATGGTTGTACCAACCGGGGTAATTAGCCTCGAACCACTCCA
TCGCCAGGCTATCCGGGAGTTCCATGCGGATGCCGGTCAGCGGCCACAGGGC
AAGGGCCAGCAGCGCGAGATCGTGGTGTGCCCAAACGGCGTCTTTCTTGGCG
TCGGGCAGGCACTTTGGCGTTTTGACGCCGAACTGCTGCAGTCGGCCCAGCC
AGATGCCAGCCC

REFERENCES

Aleklett, K., and C.J. Campbell. 2003. The Peak and Decline of World Oil and Gas Production. Uppsala University: Sweden. Association for the study of Peak Oil & Gas. http://www.peakoil.net.

Archer, D.E. 2007. Methane hydrate stability and anthropogenic climate change. Biogeosciences 4:993–1057.

Arp, D. J. 1999. Butane metabolism by butane-grown 'Pseudomonas butanovora'. Microbiology 145:1173–1180.

Baldwin, B.R., C.H. Nakatsu, and L. Nies. 2003. Detection and enumeration of aromatic oxygenase genes by multiplex and real-time PCR. Applied Environmental Microbiology. 69: 3350-3358.

Brisbane, P.G., J.N. Ladd. 1965. The role of micro-organisms in Petroleum exploration. Annual Review Microbiology. 19:351-364.

Burgmann, H., M. Pesaro, F. Widmer, and J. Zeyer. 2001. A strategy for optimizing quality and quantity of DNA extracted from soil. J. Microbiol. Methods. 45:7-20.

Colby, J., and H. Dalton. 1976. Some properties of a soluble methane monooxygenase from Methylococcus capsulatus strain (Bath). Biochem. J. 157:494-497.

Hamamura, N., R.T. Storfa, L. Semprini, and D.J. Arp. 1999. Diversity in butane monooxygenases among butane-grown bacteria. Appl. Environ. Microbiol. 65:4586–4593.

Han, D.D., R. Chen, E.R. Hill, M.R. Tilley, and H.H. Gu. 2006. Cause and solutions to the polymerase chain reaction smear problem in genotyping. Anal. Biochem. 353:296-298.

Higgins, I.J., D.J. Best, R.C. Hammond, and D.Scott. 1981. Methaneoxidizing microorganisms. Microbiol. Rev. 45:556–590.

Hirsch, R.L., R. Bezdek, and R. Wendling. 2006. Peaking of world oil production and its mitigation. AIChE J. 52 (1) 2-8.

Hitzman, D.C., B.A. Rountree, J.D. Tucker, and S. Smith. 2002. Integrated microbial and 3-D seismic Surveys discover Park Springs (Conglomerate) field and track microseepage reduction, in Surface exploration case histories: Applications of geochemistry, magnetics, and remote sensing, D. Schumacher and L.A. LeSchack, eds., AAPG Studies in Geology No. 48 and SEG Geophysical References Series. No. 11, p. 59-65.

Kinnaman, F. S., D.L. Valentine, and S.C. Tyler. 2007. Carbon and hydrogen isotope fractionation associated with the aerobic microbial oxidation of methane, ethane, propane and butane. Geochim. Cosmochim. Acta 71, 271-283.

Kotani, T. T., Yamamoto, H. Yurimoto, Y. Sakai, and N. Kato. 2003. Novel acetone metabolism in a propane-utilizing bacterium, *Gordonia* sp. strain TY-5. J. Bacteriol. 185:7120-7128.

Kotani, T., H. Yurimoto, N. Kato, and Y. Sakai. 2007. Novel acetone metabolism in a propane-utilizing bacterium, *Gordonia* sp. strain TY-5. J. Bacteriol. 189:886–893.

McCain Jr., W.D. 1990. The Properties of Petroleum Fluids, p.1-2, 2nd ed. Pennwell Books. Tulsa.

McDonald, I.R., E.M. Kenna, and J. C. Murrell. 1995. Detection of methanotrophic bacteria in environmental samples with the PCR. Appl. Environ. Microbiol. 61:116–121.

Ross K.S., N.E. Haites, and K.F. Kelly. 1990. Repeated freezing and thawing of peripheral blood and DNA in suspension: effects on DNA yield and integrity. J. Med. Genet. 27:569–70.

Sagova-Mareckova, M., L. Cermak, J. Novotna, K. Plhackova, J. Forstova, and J. Kopecky. 2008. Innovative methods for soil DNA purification tested in soils with widely differing characteristics. Appl. Environ. Microbiol. 74:2902–2907.

Sluis M.K., L.A. Sayavedra Soto, D.J. Arp. 2002. Molecular analysis of the soluble butane monooxygenase from 'Pseudomonas butanovora'. Microbiology. 148:3617-3629.

Stienkamp, R., W. Zimmer, and H. Papen. 2001. Improved method for detection of methanotrophic bacteria in forest soils by PCR. Current Microbiology. 42:316-322.

Stolyar, S., S. Van Dien, K.L. Hillesland, N. Pinel, T.J. Lie, J.A. Leigh, and D.A. Stahl. 2007. Metabolic modeling of a mutualistic microbial community. Mol. Syst. Biol. 3: 92.

Streifel, Shane. 2006. "Impact of China and India on Global Commodity Markets: Focus On Metals and Minerals and Petroleum." *Background paper for Dancing with Giants: China, India, and the Global Economy.* Institute for Policy Studies and the World Bank, Washinton DC.

Theisen, A.R. and J.C. Murrell. 2005. Facultative methanotrophs revisited. J. Bacteriol. 187:4303–4305.

van Beilen, J.B., Z. Li, W.A. Duetz, T.H.M. Smits, and B. Witholt. 2003. Diversity of alkane hydroxylase systems in the environment. Oil Gas Sci. Technol. 58:427-440.

Wagner, M., J. Piske, and R. Smit. 2002. Case histories of microbial prospection for oil and gas, onshore and offshore in northwest Europe, in Surface exploration case histories: Applications of geochemistry, magnetics, and remote sensing, D. Schumacher and L. A. LeSchack, eds., AAPG Studies in Geology No. 48 and SEG Geophysical References Series No. 11: 453–479.