SETG, a Search for Extraterrestrial Genomes: An *in situ* PCR Detector for Life on Mars Ancestrally Related to Life on Earth

Gary Ruvkun, Professor of Genetics, Harvard Medical School
Michael Finney, Former Chief Scientific Officer and Principal, MJ Research
Maria T. Zuber, E.A. Griswold Professor of Geophysics and Planetary Science, MIT
Chris Carr, PhD
George M. Church, Professor of Genetics and Director of the Lipper Center for Computational Genomics, Harvard Medical School
Walter Gilbert, University Professor, Dept. of Molecular and Cellular Biology, Harvard University
Stephen Quake, Associate Professor of Applied Physics, Caltech
William F. Mayer, Associate Director, Center for Space Research, MIT

Proposal Summary

On Earth, very simple but powerful methods to detect life by the DNA polymerase chain reaction (PCR) are now standardly used. A soil sample from an extreme environment can be surveyed for the signature of life, a DNA fragment of a gene that is universal to life on Earth in less than 2 hours in any standard molecular biology laboratory. Due to massive meteoritic exchange between Earth and Mars (as well as other planets), a reasonable case can be made for life on Mars or other planets to be related to life on Earth. In this case, the supremely sensitive technologies used to study the extremes of life on Earth can be applied to the search for life on other planets. We propose to develop a PCR detector for in situ analysis on other planets, most immediately, Mars. This instrument is so sensitive it should allow the detection very low levels of microbial life on Mars, and will determine its phylogenetic position by analysis of the DNA sequence of the genes detected *in situ*. Our team includes leaders in gene analysis from Harvard University plus the management and engineering staff of MJ Research, a 400 person company that is the second largest manufacturer of PCR machines in the world, plus the planetary science and space engineering expertise of the MIT team members. We propose to develop a prototype of the DNA amplification module of this instrument which will be validated using terrestrial samples. We are aiming for an instrument that is less than 5 Kg and requires less than 10 Watts during any experimental run. We also propose to explore contamination reduction protocols and to expand the DNA probes used to explore the boundaries of detectable life on Earth, to maximize our chances of detecting life that is divergent from life on Earth.

Introduction

Strategies for detecting life on other planets have sought to avoid the assumption it would share any particular features with life on Earth. The most general strategies — seeking informational polymers, structures of biogenic origin, or chemical or isotopic signatures of enzymatic processes — look for features that all life is expected to exhibit. This generality comes at a cost: the strategies are not particularly sensitive, and more importantly, there are abiological routes to these life signatures. However, if life on Earth is actually related to life on other planets, we can use a far more powerful and information-rich technique developed to detect the most extreme forms of life on Earth.

Increasing evidence, such as the low temperature transfer of ALH84001, and theoretical calculations suggest that objects capable of carrying life have been transferred between solar system bodies with significant frequency. In addition, extremophiles have been discovered in Earth environments with high radiation and frozen conditions which, while not as extreme as those on Mars and other planets, demonstrate the incredible adaptability of microbes and suggest that habitable zones are much broader than previously thought. Together these facts raise the possibility that life could have been transferred between Earth and Mars perhaps early in the history of the solar system, and could survive on Mars to the present day. Therefore, we propose to build a very low power and lightweight instrument to test for life on other bodies, most immediately Mars, using the most sensitive known detector for Earthly life.

16S RNA gene PCR, the most sensitive detector for life on Earth

This detector is an amplification strategy called the Polymerase Chain Reaction (PCR) that is based on artificial replication of DNA. PCR is a technique which is used to amplify the number of copies of a specific region of DNA, in order to produce enough DNA to be further analyzed. In order to use PCR, one must know the exact DNA sequences which lie on either side of a given region of interest in DNA. One need not know the DNA sequence in-between. A DNA sequence is the precise order of appearance of 4 different deoxyribonucleotides. The 4 components are: Adenine, Thymidine, Cytosine and Guanine, abbreviated A, T, C and G, respectively. The arrangement of this 4-letter alphabet is the DNA sequence.

The PCR strategy for life detection emerged from the exploration of the diversity of life, which revealed about 500 "universal genes" that are carried in the DNA of every known living thing on Earth. The gene that has changed the least over the past 3-4 billion years is the 16S (or the related eukaryotic 18S) ribosomal RNA gene. Ribosomal RNAs are the main structural and catalytic components of the ribosome, a molecular machine that translates RNA into proteins.

It is the slow rate of change of the 16S gene that makes it the best detector of life. Within the ~1500 nucleotides of the 16S gene, there are multiple 15 to 20 nucleotide segments that are exactly the same in all known organisms. These regions of the 16S gene are essential for its catalytic activity and have remained unchanged over billions of years.

The technology of PCR involves adding stable 15-20 nucleotide long DNA primers, a stable enzyme nucleotide triphosphate monomers, and a simple heat pump that thermally cycles 20-30 times in 2 hours. To amplify 16S genes from a crude sample, universal DNA primers from the ribosomal RNA gene that are about 18 bases long, oriented towards each other, and about 1000 bases apart are added to crudely purified DNA isolated from an environmental sample (for example, 1 ml of sea water or 1 gram of earth). For the ribosomal genes, the DNA primer 5' GTGCCAGCAGCCGCGGTAA 3' which corresponds to nucleotides 515 to 533 of a ribosomal gene, and 3' TTCAGCATTGTTCCAWYGGCAT 5' which corresponds to the base pairing complement of nucleotides 1492 to 1510 are added to an extract prepared from soil (M, Y, and W are codes for mixtures of two such nucleotides necessary to capture all 16S genes). Upon heating to 95°C and then cooling to 55°C, these DNA primers pair with their complement on each DNA strand, even if there are only a few DNA molecules in a sample. After heating to 75°C, the DNA polymerase will polymerize the nucleotide monomer components also in the tube to duplicate the DNA strands. There will now be four strands, where originally there were only two. If one repeats the thermal cycle with all the same components in the same tube, now there will be eight strands; repeat again - now 16, etc. Thirty cycles will produce one billion (2^{30}) copies of the original sequences. Because the DNA polymerase enzyme used derives from a thermophilic microbe, it can survive repeated cycles of heating to 95°C. The amplified DNAs from the PCR can be analyzed for size or DNA sequence. PCR will even amplify complex mixtures of 16S ribosomal RNA genes from communities of organisms in environmental samples. Thus, **PCR with DNA primers corresponding to the conserved elements can be used to amplify DNA from any species more than a billion fold, without need to isolate, culture, or grow the organism in any way (9).**

An example of two universal DNA primers amplifying a bacterial 16S gene

```
5'      GTGCCAGCMGCCGCGGTAA  515F primer  ------->
5'AACTCTGTGCCAGCAGCCGCGGTAATTCCAGCT------------------GTGAAGTCGTAACAAGGTAACCGTAGGGGA
3'TTGAGACACGGTCGTCGGCGCCATTAAGGTCGA------------------CACTTCAGCATTGTTCCATTGGCATCCCCT
                                        <-----        TTCAGCATTGTTCCAWYGGCAT 5'
                                           1512 universal primer
```

The PCR approach has added advantages of extreme sensitivity and robustness. PCR can detect a single DNA double helix in a crude sample**.** The biochemical processing of the sample can be as crude as a cheek swab from humans to agitation of dirt for soil microbes.

The DNA sequence that is amplified between those primers is key to the analysis. The DNA sequence between the universal 16S gene primers contains so much information that organisms detected only by their 16S gene sequences are routinely classified based only on that information. This DNA sequence carries information about the organism from which the ribosomal RNA gene is derived, and can allow a new organism to be fit into the tree of life. Thus the detected product is a unique biosignature.

Currently, hundreds of research groups use 16S ribosomal RNA PCR primers to prospect for new archaeal and bacterial species from a wide range of environments. Most of the life that is detected by PCR cannot be grown in the lab, suggesting either very slow growth rates or very particular growth conditions not met in the lab. Thus the previous culture-based exploration of microbial diversity missed 99% of the living world. Such surveys of extreme environments have expanded habitable zones from below 0 °C to over 110 °C, from acidic hot springs to highly radioactive reactor pools, to deep in the crust of the Earth, and has allowed particular 16S gene sequences to be assigned to particular metabolic strategies.

PCR is simple, lightweight, and low power

PCR instruments involve simple and efficient heating and cooling of 0.1 to 100 microliter samples. The technology is very mature, with thousands of thermal cycling machines installed in small labs all over the world, and field PCR thermal cyclers used, for example, in the military to detect biological warfare agents. They are as standard in the modern molecular biology laboratory as toasters are in kitchens. A typical small thermal cycler not optimized for space flight weighs 3 kg and uses 100W. Only tiny amounts of energy are actually needed to cyclically heat and cool the 10-100 microliters of fluid in a typical PCR reaction, and to detect the product of that amplification. We estimate that the fluid handling components to handle the soil samples and transit to the PCR module could weigh less than 100g. The thermal cycler for amplification could weigh just a few hundred grams. The DNA sequence analysis chip and CCD electronics to read the chip scanner could be similar in size to a portable CD player.

The case for life on Mars to be ancestrally related to Life on Earth

Life on Mars should resemble life on Earth if there has been exchange of living cells between Mars and Earth. For this to be the case, organisms must have survived the passage, and adapted to a new environment. The ability of a crustal sample to escape Mars and land on Earth without sterilizing heating is demonstrated by the low temperature transfer of ALH84001. Experimental observations of 4 Km/sec test impacts favor this observation: non-shocked and therefore essentially unheated high velocity ejecta occur. It is thought that in any impact, the reflected shock wave interferes with the impact shock wave to allow a region of no shock heating that nevertheless accelerates to escape velocity. Theory also supports this view: for 1 to 20 Km impactors, the low temperature meteoritic exchange between Mars and Earth has been estimated at 10^8 to 10^{11} ejecta with an average diameter of 0.3 to 6 meters. Because of cosmic and isotopic decay radiation, transit times are critical. About 3 to 5% of the ejecta from Mars arrive on Earth within 10 million years, but perhaps some in as short as decades. Transfer from Earth to Mars is about 100-fold less efficient, but is still substantial.

In transit between Mars and Earth, radiation and desiccation are the major challenges to survival. Desiccation in the vacuum may be less of a problem if the organisms are "sealed" in a salt crystal or rock fissures from gas exchange with the vacuum. Shielding against UV radiation is afforded in just a few mm of rock. Survival of spore forming Bacillus subtilis bacteria in the vacuum, temperature extremes, and high radiation environment of space for 6 years (30% survival within salt crystals), and in one report of survival of Streptococcus for 3 years on the surface of the moon suggest that microbes can survive outside the atmosphere for periods of a year. UV radiation damage appears to be easily shielded, for example, by simple embedding in rock, though survival for millions of years is difficult to estimate. The ejecta size is likely to be important for radiation shielding; a 6-meter ejecta provides sufficient shielding against cosmic radiation for about 100,000,000 years of viability (to 10^{-6} of the initial innoculum) whereas 0.3 meters affords about a million years.

After transit between planets, atmospheric reentry heats the surface of a meteorite to beyond the survival temperatures of any known extremophile. However, because the transit of the atmosphere at 15 Km/sec takes just seconds, and because the black body temperature of a meteorite, for example at Mars is 220 °K, and because of spallation of the burning surface crust, the temperature inside the reentering meteorite can be below 100 °C, consistent with the observations of ALH84001. Such low heating reentries may be the rule; in the few cases where meteorite falls have been retrieved instantaneously, they have been cold to the touch.

Thus, calculations of collision dynamics, meteor impact rates and celestial mechanics are compatible with a model of exchange of viable microbes between Mars and Earth.

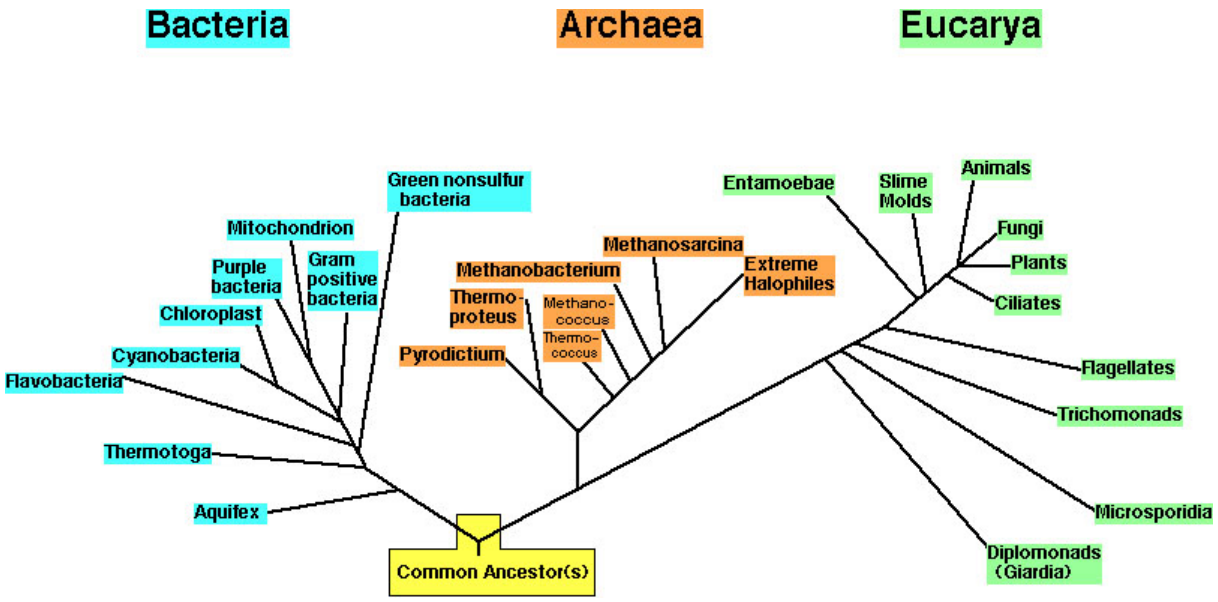
The distinct environments of the two planets at present might not allow an organism adapted to one planet to grow on the other. But meteoritic exchange in the solar system was 100 to 1000x more intense during the heavy bombardment stage 4 billion years ago. There are signs of numerous fluid flows a possible ancient ocean, and sedimentary formations on Mars that suggest a warmer and wetter Mars 3 to 4 billion years ago, an environment more similar to Archean Earth.

Fossil evidence suggests that by the Archean period, microbial evolution on Earth had already proceeded to the point of modern microbial morphologies, and biosignatures suggest that enzymatic carbon metabolism with isotopic fractionation had evolved by then. Because all known organisms have one or more copies of the 16S ribosomal gene, all organisms are thought have inherited their ribosomal RNA gene from a common ancestor. This common ancestor has been hypothesized to be an archaeal-like hyperthermophile 3 to 4 billion years ago whose metabolism exploited oxidation/reduction gradients. Thus at the time of maximal meteoritic exchange 3.5-4 billion years ago, microbial life on Earth may have already possessed a shared core of 500 genes, including the 16S ribosomal RNA gene. The last common ancestor with life on Mars may have also shared this core of genes. **Thus at**

the point of high meteoritic exchange, there may have been microbial life on Earth detectable by 16S gene PCR and an environment on Mars more similar to Earth than today.

Subsequently, the environments on Mars and Earth have diverged: the appearance of oxygen on the Earth 2 billion years ago led to decreased UV radiation, while Mars lost its atmosphere as its magnetic field decayed, causing an increase in UV flux, cooling of the surface, and loss of surface water. Current life on Mars would need to survive temperatures and pressures below the triple point of water, high UV flux and the oxidizing surface chemistry induced by UV radiation. The UV flux on Mars is now about 1000x that on the Earth. But only a thin layer of soil is sufficient to protect microbes from this UV flux. In addition, the redox gradient resulting from UV irradiation might actually power microbial metabolism, just as redox gradients in the Earth's crust drive chemolithotrophic metabolism.

While there is no doubt that Mars is currently an extreme environment, given the adaptability of microbial life on Earth, it is not unreasonable to propose that Martian microbes could have adapted to the gradual decline in water, temperature, and UV protection over the past few billion years. **And just as the adapted and diverged microbes on Earth still bear the signature of their common ancestry in their 16S ribosomal RNA genes, the Martian biota may also bear this signature.**



But any Martian biota has been geographically as well as genetically more isolated from the Earth biota since the dramatic decline in meteoritic exchange 3.5 billion years ago, and thus is likely to be more deeply branching in the phylogenetic tree of 16S RNA sequences. Even though there may continue to be exchange, Mars is more likely to be more isolated than any possible Earth ecosystem. **Therefore, we expect that any 16S gene detected on Mars will branch deeply compared to the phylogeny of 16S RNA gene sequences of known Earth life. Detection of such a deeply branching 16S gene would be a compelling signature of life on Mars.**

Even if life did transfer between Mars and Earth 4 billion years ago and thrive in the early Mars environment, it may now only thrive in very particular refugia, for example deep in the crust where the temperature rises above that of the frozen surface or at particular volcanic thermal vents. There may be regions on Mars where liquid water is not in short supply, for example near the polar ice caps. In addition, the D/H ratio of water in Martian meteorites suggests a much larger reservoir of water in the crust that is not in equilibrium with the atmosphere. This water is predicted to be liquid a few kilometers into the crust, where temperatures rise above surface temperature. In addition, there is evidence for recent Martian volcanism, suggesting sources of temperature gradients and fluid flows just below the Martian surface near these sites. There may be local hydrothermal systems near the regions of recent volcanism. However, the negative (albeit insensitive) Viking seismic data and the build up of massive volcanic cones do not favor active tectonics, a key element for hydrothermal vents and the biology that thrives in their energy and nutrient flows.

In 2007 or any of the missions in the foreseeable future, we cannot expect any polar regions or particular oases to be technically within reach; landings will be constrained by navigational limitations and celestial mechanics to obstacle free zones and equatorial positions. But PCR, unlike other detection strategies, could detect the dispersal of a few microbes or microbial genomes from those refugia. **Even if microbes only flourish today in particular oases on Mars, PCR should be able to detect any that are dispersed by windstorms.**

It is also possible that life did transfer between Mars and Earth but is now extinct. We do not expect to detect microbial fossils using PCR; it has been used to detect DNA in samples thousands of years old but not

millions of years old (though viable halophilic bacteria have been reported in 250 million year old salt crystals, but this has been debated).

Why do the PCR on Mars?

The most obvious use of this PCR analysis is on any Martian sample returned to Earth. But the issues of contamination of any returned Martian sample are even more profound than contamination of equipment sent to Mars. Earth is simply covered with microbial life. And while a returned sample can be assayed more completely, it will come from a single point on the surface, selected for ease of landing and return, with all the limitations imposed by celestial mechanics and topography considerations. In situ analysis with very light and low technology PCR detectors could target multiple locations, escaping the jackpot effects of single site selection. Due to Jupiter's gravitation, return missions will be even more unlikely from Europa. **The low-mass, low-power PCR detector we propose to develop could be part of a more extensive life detection package for multiple planetary missions.**

Planetary protection via the SETG PCR detector

While we propose the development of this instrument with the goal of detecting life on the Mars 2007 mission, the success of the experiment will depend on the landing site. Currently, safety, navigation, and orbital dynamics can take precedence over the selection of the most promising life refugia landing sites. If we detect no amplification of 16S RNA genes at this landing site, constrained by technical limitations on navigation and craft design, it would **set a baseline for future Mars exploration**. Future missions, after the development of global Mars communications system, will be capable of targeting landings to regions of possible water flows and sediments. These are more likely to harbor current life. At each landing site, we propose to use the PCR approach to seek life, comparing to that baseline.

It is also possible that we may not be able to purge the landing craft of Earthly life, and that we will detect a 16S signal that we recognize. PCR analysis of landing sites would set a baseline for how much contamination is actually brought from Earth to Mars and what sorts of microbes make the transit. Thus, the PCR detector does not only seek extraterrestrial life, it also constitutes an element of a planetary protection detector.

PCR vs. other life detection strategies

Competing life detection technologies detect biosignatures, such as lipids, proteins, or nucleic acids, with a sensitivity of a few thousand monomer subunits polymer per gram of input soil. But there are abiogenic mechanisms for the synthesis of amino acids, including meteoritic deposition. Proposals to search for enantiomeric organic molecules as evidence of life are also haunted by possible abiological sources of asymmetry. It is doubtful that isotopic fractionation observations could be so sensitive, though they do have the advantage of being able to observe fossil biosignatures.

No other life detection technology can rival PCR for sensitivity. It can detect a single DNA strand in a pure preparation. Rich environments on Earth such as soil in temperate zones have 10^8-10^9 microbes per gram of soil. Samples of rock or deep core drillings typically have “only” 10^4 microbes per gram (N. Pace personal communication). Our goal during this phase of instrument design is to be able to detect with a sensitivity of 100 microbes per gram of input soil. We cannot predict what sorts of microbial densities could be found on Mars: the energy flux at the surface would allow the sorts of densities found on Earth, but the failure to detect organic molecules on the Martian surface suggests low microbial densities.

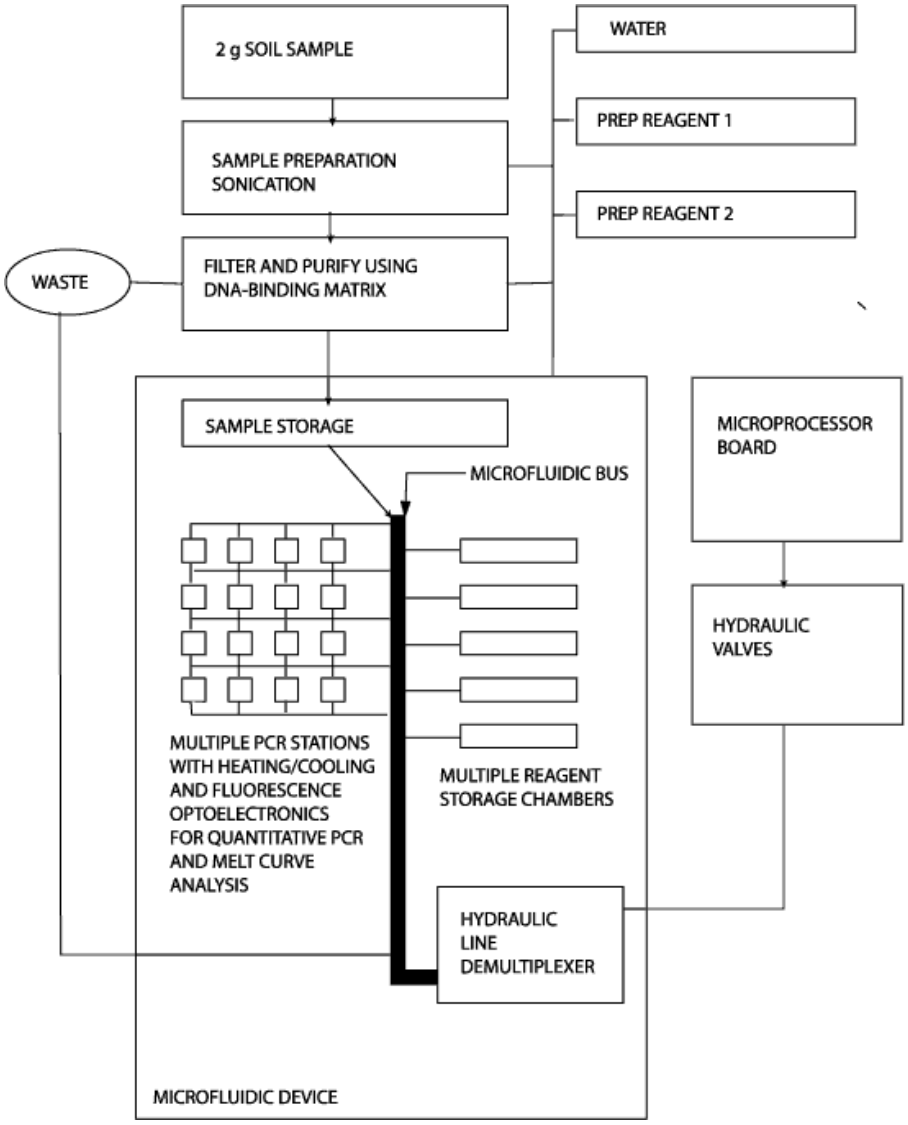
Of course, the hypersensitivity of PCR makes it prone to contamination problems. But the sensitivity can be tuned: more cycles of amplification make the technique more sensitive, and more prone to artifact and contamination detection.

But it is the specificity of the PCR signal that is also important. A positive signal with the DNA sequence features of the ribosomal RNA genes can only come from life. There is no abiological route to the signal. We must be concerned about contamination, but the **prediction that a 16S RNA gene indigenous to Mars will branch more deeply is quite specific**.

Overview of the proposed instrument

We propose the engineering design and a breadboard model of the thermal cycler and associated fluidics, and of the double-stranded DNA detector and melting point analysis module. We will also design and integrate into a single instrument, the sonicator to disrupt the soil sample and a DNA extraction module upstream of the PCR element.

First, we describe our best current thinking about the *in situ* analysis protocol. Second, we describe currently-favored technical approaches to performing that protocol. Third, we discuss issues of sensitivity, contamination minimization and detection.



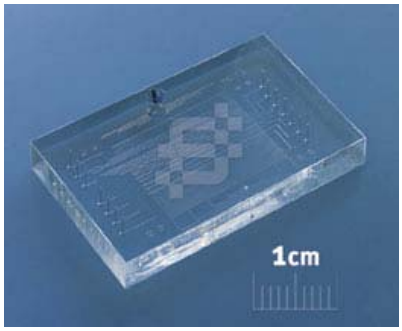
The analysis protocol: **Analysis will be performed on a 2g soil sample sieved to 500 microns maximum size. We assume that any landing craft will have a soil retrieval and sieving element that will deliver a few grams to our DNA isolation module. All fluid handling and sample movement will take place in a pressurized line. Approximately 5 milliliters of a pH buffer solution containing reagents to chemically protect DNA (reducing agents, free-radical scavengers, and metal ion chelators) and absorb PCR inhibitors commonly found in soil will be added. The mixture will be sonicated using a simple solid state low power off-the-shelf component. The combination of osmotic shock and sonication is expected to break open any microbes present, but a concentrated detergent/protease solution will be mixed into the sample post-sonication to ensure complete lysis.**

After settling, 3 milliliters of the supernatant will be decanted and filtered to remove particles larger than 0.5 microns. The filtered lysate will be passed slowly over a DNA-binding matrix. These matrices purify DNA from complex mixtures and are commonly used and commercially available from companies such as Qiagen (Valencia, CA). The matrix will be washed with 3 milliliters of concentrated salt solution twice to remove any chemicals that do not bind tightly the matrix. Any DNA bound to the matrix will be eluted in 50 microliters of a dilute buffered and chelated salt solution. DNA in such a solution is stable for years. This 50 microliter stock will serve as the raw material for subsequent experiments. Current thinking is that ten experiments will be performed, each using approximately 5 microliters of the sample.

Microfluidic valves and tubing

Downstream of the DNA isolation module, will be a microfluidic sample handling valving and mixing station. Stephen Quake, an Associate Professor of Applied Physics at Caltech and a group of engineers and scientists at Fluidigm, a biotech company that specializes in microfluidic design and manufacture will be closely involved in the design of this microfluidic module.

The valves on this microfluidic chip use pressurized gas lines to open and close fluid paths with pressurized control lines. A cascade of these pressurized lines can also pump fluids between chambers or in circuits, for example during PCR. They have already created microfluidic chips which can move fluids between dozens of mixing and analysis modules on a monolithic matrix that is less than 1 cm square and weighs just a few grams. The material used is polydimethylsiloxane, a soft rubber that is commonly used in industrial sealants and adhesives. This matrix is now being used for many microfluidic applications, including PCR (ref: Liu, J. Enzelberger M, Quake, S. A nanoliter rotary device for the polymerase chain reaction. Electrophoresis 23: 1531-36 2002). The material may need further space validation (I tried to find out about this material on the Huntsville web site and found it listed but could not figure out if it actually space or mil spec approved). **This sample handling microfluidics has low DNA binding characteristics and will be used for the initial 16S PCR as well as the downstream query oligo PCR.**



Shown is a microfluidic chip from the Fluidigm web site. This rubber like material is filled with 20 micron sized sample and control paths that are connected to pressure solenoids and fluid sources.

Each PCR experiment will be set up by mixing 5 microliters of sample with 5 microliters of a reaction mixture concentrate. Different concentrates will be available, containing different DNA primers and thus capable of directing different amplification reactions.

Thermal cycler: **Thermal cycling will be achieved by moving the sample through a spiral or circular path in the microfluidic matrix, where three sides of the spiral cylinder are maintained at three different temperatures (see Figure). The components of the PCR reaction will be mixed with the DNA from the Qiagen column by laminar movement in the path. The actual DNA amplification will be achieved by serial heating and cooling by pumping a bolus of the mixture of Martian DNA, ribosomal RNA gene oligonucleotides, and PCR reaction components along a microfluidic path, between heating elements at 95 °C to denature DNA, 55 °C to allow hybridization, and 75 °C for DNA polymerase extension. For each turn of the spiral, the sample will experience a single thermal cycle of the three temperatures in series. By avoiding active temperature changes, power consumption and failure modes are minimized. LED/photodiode pairs will provide feedback to ensure that the sample spends the proper amount of time in each temperature zone.**

The time at each temperature will be determined by the pressure difference across the capillary tube that drives movement, the resistance of the microfluidic path, and the length of the path along the temperature block. The thermal cycling will be achieved by flow along the fluid path, transiting between the heat blocks up to 40 times as the fluidics winds around the cylindrical heating elements. Sensors for the meniscus will give feedback to the pressure valves to control the movement.

The heat blocks will be maintained at a particular temperature, with negligible heat loading from the sample of only 10 to 20 µl; the major energy drain will be heat loss to the instrument chamber, but in the ambient temperature of 10-20 °C, which is necessary to keep the water in the experiment liquid, and atmospheric pressure of 6 mbar in the instrument outside the pressurized tubing, the heat flow from the elements to the chamber will be minimal.

A disadvantage of this approach is that the surface area of the thermal cycling reaction chamber is very high compared to the usual 10 µl nearly spherical cycle chamber. This can decrease detection efficiency if the DNA at the early stages of the amplification bind to the surface, especially at the denaturation step of the PCR. We will optimize the sensitivity of the instrument by exploring surface types and surfaces coatings to minimize DNA and other reagent binding. We are also exploring an alternative format with much lower surface area: moving the sample cyclically between 3 different temperature stations by changes in air pressure. If we find that surface binding is a serious problem, we will instead use a small surface area very low heat capacity (low mass) reaction chamber and thermal heating and cooling to amplify.

We expect that the microfluidic design will allow multiple parallel PCR paths to allow the soil sample to be tested with multiple primers. Thus we will have multiple opportunities to amplify DNA in those samples.

We outline the 4 methods we will use to detect amplified 16S gene DNA below, but stress that the weight limitations of the rover will constrain us to a few dozen classification PCR analyses. In a future lander mission, where more weight is afforded a next generation of this instrument, we could include the most desirable DNA sequence determination module.

Detection of amplified 16S genes

We will use four methods to give increasing levels of detailed characterization. The methods are increasingly good at discriminating the identity of the DNA. First, the amount of DNA present at each cycle will be assayed by a fluorescent probe that detects double stranded DNA. This measurement can be used to calculate the number of molecules in the input sample. Second, the DNA will be analyzed by melting curve to determine an approximate size and composition. Third, the amplified product will be re-amplified with a different set of DNA primers to query it for other universal sequence features of the 16S gene. Finally, the DNA sequence will be interrogated using a biochip-type optical detector to detect many other DNA sequence features.

a. Real time PCR detector

The most basic type of analysis will be to detect the amount of DNA produced at each cycle in the PCR amplification. The dye SYBR Green (Molecular Probes, Eugene OR) fluoresces brightly only when bound to double stranded DNA. Fluorescence can be detected at specific points in the PCR by blue LED excitation at 490 nm and filter/photodiode detection at 525 nm. By detecting the amount of DNA produced at various points along the spiral path, that correspond to particular PCR cycles, it is possible to back-calculate the amount of DNA present in the initial reaction mixture.

b. Melt curve analysis

A blue LED/filter/photodiode detector on the 95 °C heating element can detect the melting of a DNA double helix by loss of SYBR Green fluorescence. A plot of DNA melting vs time or temperature contains information on the number of independent melting domains in the product and their relative lengths and nucleotide content. These profiles can be compared with those for the expected product of a PCR, and to eliminate artifacts such as primer dimer formation. The additional hardware necessary for this test is an emitter-detector pair and a slowly changing temperature block, for example one of the heating elements when not being used for thermal cycling.

c. Secondary PCR analysis

A method of directly confirming the identity of a product is to use a small quantity of it as input for a second PCR, using one or two primers complementary to a sequence in the interior of the first product. The second PCR product is predicted to be shorter than the first product by a particular amount, and amplified in a smaller number of cycles, related to the dilution factor. For example a set of DNA oligonucleotides that also universally detect ribosomal genes (906-922F=GAAACTTAAAKGAATTG and 1407-1391R= GACGGGCGGTGWGTRCA, where K = G or T, W = A or T, and R=G or A) but prime within the 519 to 1492 region amplified in the first step will be used. In this way, the sequence of the amplified DNA is queried with these two distinct, nested but also universal DNA oligonucleotides. Bonafide 16S gene segments from the first round of PCR diluted one million-fold will amplify in 20 cycles in the second round of PCR, and generate a DNA fragment of about 500 bp for all known organisms, whereas artifactual DNA segments will not. This same query DNA oligonucleotide protocol could be repeated with other primers corresponding to universally conserved (or conserved in particular clades of microbes) oligonucleotides.

An additional advantage of the re-analysis of any DNA fragment is that the second PCR amplification can be used to increase sensitivity after the first amplification**. Such nested PCR reactions are commonly done in the lab. In this way, the signal is amplified by two different chain reactions. This additional sensitivity may be crucial for the detection of low titers of life in the 2-gram sample of soil tested.**

We expect to include 20-100 oligonucleotides from the 16S gene in the instrument, plus 5-10 oligonucleotides for a generic DNA detector also described below. The hardware required for this secondary amplification is principally more microfluidics for moving and mixing liquids. Because the microfluidic module is monolithic and lightweight, it adds very little weight to have this analysis capability. The analytical oligonucleotides can be stored as dried DNA in one of hundreds of plumbed reservoirs in the microfluidic path. We intend to design the optimal 100 classification oligonucleotides during the first two years of this grant. The Church lab is a world leader in genome analysis and will assist in the informatic analysis.

d. DNA sequence determination

The definitive analysis of any PCR product is a DNA sequence determination. The current laboratory DNA sequencing technique, high-resolution electrophoresis, is not practical on Mars. It is possible, however, to determine the DNA sequence using a "biochip" containing a two-dimensional array of thousands of DNA oligonucleotides that can each potentially base pair with a distinct region of a fluorescently labeled 16S gene PCR product. Each 20 micron element on the gene array has about a million copies of a distinct 8 to 25 nucleotide long DNA segment adhered to that position. If any of those DNA segments are complementary to a region on the 1000 base 16S gene DNA fragment (fluorescently labeled by covalent linkage to one of the initial 16S PCR primers, standard technology in molecular biology), it will base pair to that position on the gene array, and that position will now be fluorescent. Excitation can be accomplished with a blue LED array. Detection can be accomplished by a CCD camera with appropriate filters. A prototype 16S gene array using 20-nucleotide query elements has already been developed by Gary Anderson and Wendy Wilson at Lawrence Livermore Laboratory for the detection of microbial diversity in natural systems. They are excited about collaborating with us. But we do not propose to develop this module for the initial instrument design.

Liquids: **Liquid samples must be confined in a pressurized system at all times, as water will boil at Martian atmospheric pressure. Some phases of the protocol involve heating the liquid to 95 °C, so a pressure of about 1 bar will be necessary in at least parts of the system. In addition, liquid samples need to be moved from station to station within the instrument. To address both of these problems, we will maintain the fluid system at pressure of 1-3 bar in a combination of tanks, tubes, and monolithic microfluidic devices. Samples will be moved within the system by selectively venting certain points to the Martian atmosphere. Pressure will be supplied by a small tank of liquefied non-explosive gas which will provide roughly 4 bars at 5 °C. We will develop these gases and protocols during at MJ Research and the Center for Space Research.**

Approximately 50 ml (50 g) of water will be transported as ice in a tightly-sealed vessel. On Mars, the ice in the vessel will be melted and pressurized to serve as the source of all liquids. Solutions will be prepared by admitting water to prepared chambers containing the dried solution components. Liquid movement monitoring and volume measurement will be accomplished by LED/photodiode pairs that can detect the presence of absence of fluid in a channel by means of change in index of refraction.

Biochemical reagents and the space environment

For the flight instrument, all of the chemical and biochemical reagents will be stored desiccated. Many of the components will be stored within the microfluidic chip dried until they are hydrated by the valving of water to that position. The total weight of these dried reagents will total less than 10 grams. Most or all of the components are commonly stored at –20 °C in a desiccated format. They are known to be stable on Earth to cold and low level of radiation, but their stability in vacuum and high solar radiation flux will need to be addressed during this development phase. The Taq DNA polymerase is the only enzyme needed for this analysis. It can be stored as a desiccated trehalose protected tablet at room temperature. We will test its survival of mock interplanetary cruise and Mars surface temperatures and radiation environments.

Engineering milestones of the thermal cycler module

We propose the engineering design of the microfluidic and thermal cycling modules and the double stranded DNA detection/melt curve analysis module between January 2003 and June 2004. We will select the components and build a functioning breadboard model that will take the output of a Qiagen DNA purification column from a soil extract "doped" with 1000 copies of a particular microbe and in a spiral PCR format, amplify that DNA using 16S PCR primers. We will optimize the surface treatments and contamination reduction protocols to the point that we can achieve the sensitivity to detect 1000 copies or less of a microbial genome. We will also do the engineering design on the miniaturized version of that instrument. The Center for Space Research will perform the space engineering and validation of the instrument.

Most of the instrument development engineering will take place at MJ Research, with much instrument testing at MGH. At this point we have met as a team about a dozen times to converge on an architecture of the instrument. We have the first test instrument now and have begun protocol optimization.

We will explore surface treatments of the tubing to minimize adsorption of denatured DNA, which decreases sensitivity. We will optimize sensitivity of the instrument by treating the capillary and microfluidic

surfaces tubing with silanizing reagents and other surfactants, so that denatured DNA at early phases of the PCR do not adhere. We are currently testing a variety of surface treatments, such as silicating solutions, for adsorption of denatured DNA, as well as for increasing sensitivity in amplification of 16S ribosomal RNA genes.

Estimates of final mass and power requirements for the instrument with DNA duplex melt detector, and recursive PCR on products but without the biochip:

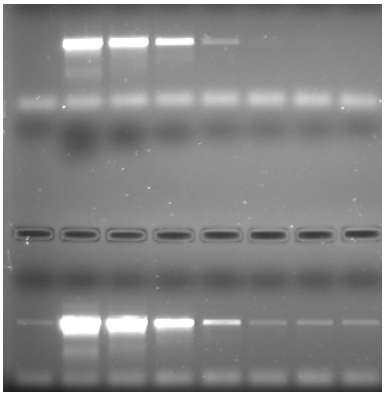
Power budget		**Weight budget = 1.5 kg**		**Volume < .3 ft^3**
PCR	20W	PCR	200g	
Electro/Mech	15W	Water	50g	
Melt liquid	5W	Tank	100g	
Sonication	5W	Gas	100g	
Daytime =	20W	Electronics	100g	
Nighttime =	3W	Valves	100g	
Peak = 33W		Fluidics	200g	
		Tubing	50g	
		Pump	200g	
		Melt	50g	
		Structure	300g	

Contamination reduction

Because PCR is so sensitive, the major issue in increasing sensitivity is decreasing false positive signals from contamination. During the development of the PCR module, the contamination issue will be confronted in the Ruvkun labs. Contamination is a major problem in terrestrial 16S PCR (45,46) due to the co-purification of bacterial DNA with the biochemical reagents used in the protocol --nucleotides, thermostable DNA polymerase, buffers and salts, and in laboratory water supplies. The DNA polymerase is cloned and expressed in E. coli, and even though it is highly purified, some bacterial DNA is present in the DNA polymerase (46). In fact, Roche sells DNA Taq Polymerase especially purified from bacterial DNA for the purpose of 16S surveys.

We will determine the baseline contamination of nucleotides and DNA oligonucleotides (45). We will develop protocols to decrease the level of microbial cell and DNA contamination of reagents. We are exploring two avenues of contamination reduction: use of micrococcal nuclease (MN), a labile calcium-dependent nuclease, and use of DNA modifying chemicals.

MN degrades nucleic acids and is inactivated completely by the addition of EGTA, a calcium chelator, as well as by the heat of PCR. We have found that we can assemble the components of a PCR reaction, including dNTPs, buffers and Taq DNA polymerase, then add calcium and micrococcal nuclease, incubate at 37 °C, and then add EGTA to inactivate the micrococcal nuclease. The EGTA solution cannot be pretreated with micrococcal nuclease, but it will be UV irradiated to destroy DNA. In this way, all of the reagents except for the oligonucleotide primers used for the PCR are treated to inactivate any microbial DNA in water, solutions, etc. We will also use micrococcal nuclease plus UV treatment to treat all of the fluid lines and the Qiagen columns before addition of DNA from a soil sample.



Shown in this figure are progressive 10 fold dilutions of archaeal genomic DNA amplified by PCR with universal 16S primers. In the bottom panel the reagents are not treated with MN, so that as the archaeal genomic DNA is diluted to less than 100 genomes, contaminating DNA from our pure water supply is amplified. In the top panel, the MN treatment destroys this DNA so that as archaeal DNA is diluted to 100 genomes in the PCR mix, no 16S gene is amplified. Thus our MN treatment removes contamination from the reagents. This is very important because many of these reagents are biochemicals processed in a manner that is not aimed at zero microbial contamination. For example, purification columns for proteins and biochemicals, as well as water purification columns are well known sites of microbial growth. While the microbes themselves can be sterilized, their DNA is not removed after such a treatment. So these cleansing procedures are critical for analysis of predicted low titer regimes expected on the surface of Mars.

To free the 16S PCR primers from contaminating DNA, we have tried synthesizing them from phosphorothioate substituted nucleotides, which are resistant to nucleases. These primers will still initiate PCR

(their 3' OH can be elongated by Taq DNA polymerase). Thus we expect to be able to add these primers to the fully prepared PCR reaction (except for the soil extract), and they will survive the MN treatment due to the S substitutions in their phosphodiester backbones. This has not worked as well as treating normal oligonucleotides with a combination of two restriction enzymes that cleave single stranded and double stranded contaminating DNA in the oligonucleotides but do not cut the oligonucleotide probes, AluI and MspI. These enzymes can be inactivated by treatment at 65 deg C. We have found that a combination of the MN treatment of everything but the oligonucleotides and AluI and MspI treatment of the oligonucleotides, followed by 65 deg heat treatment and EGTA removal of the Ca cofactor for MN effectively cleanses our reagents. PCR after these treatments is still very sensitive, but we are now establishing the limits of that sensitivity.

After these treatments, the soil sample will be added to the Qiagen column (previously treated with micrococcal nuclease followed by EGTA), eluted with elution buffer (also treated with MN), and thermal cycling begun.

We will assess the success of our anti-contamination treatments by determination of the cycle number that results in amplification of contaminant 16S genes relative to a standard addition of E coli genomic DNA. The amplification of internal control E coli 16S vs. contaminant 16S will be assessed using restriction digests of the 16S gene. If we are successful at elimination of endogenous DNA, we should find that only at very low E coli DNA dilutions do the contaminants common to water supplies emerge from our PCR analysis. This competition assay will be less subject to statistical fluctuations as DNA contamination is decreased. Our goal is to decrease contamination without sacrificing sensitivity; indeed the only real limit to sensitivity is contamination.

Our aim is to be able to detect microbial levels that are less than 100 organisms per gram of soil. We will attempt to increase sensitivity by adding DNA analogues to the soil to compete with non-specific binding of DNA (which reduces sensitivity), and act as carrier in the steps of purification and elution. For example, we may add glycogen, a polyanionic molecule, or purified tRNA as carrier to increase sensitivity. We may also find that amplification of a shorter 16S fragment, for example, 100 nt long rather than 1000 nt long, will afford better sensitivity.

The different signatures of contamination vs microbes in Martian samples

We aim to eliminate contamination from our liquids and reagents. But we may not be able to remove all vestiges of microbial life. We must therefore design controls to identify any possible Earth contamination of the reagents we propose to use on Mars. The first and most important test is the negative control. Before any Martian soil is collected, the instrument will be put through a null run where water is used as the input sample. If we are successful in our decontamination of reagents, we expect that there will be no 1 kb duplex DNA after 40 or more cycles, or even after 2 rounds of 30 cycles (nested PCR). Even if we detect contamination at, for example, cycle 40, then we may have a window from cycle 30 to 40, where we can discriminate between bona fide 16S gene signals and contamination. A true 16S signal after addition of extracts from Martian soils should appear in a number of PCR cycles that produced no signal from the null. In addition, the number of cycles required to see the amplified DNA will scale with the input of Mars soil.

The phylogenetic placement of the detected 16S gene may also argue for indigenous life on Mars. The DNA chip analysis will definitively place the amplified sequence either inside or outside the range of microbial 16S gene sequences found on Earth. If the gene sequence is more deeply branching than any known earthly clade, it would be dramatic evidence for lack of contamination.

Surveying samples from extreme environments and Mars mockup oxidizing environments

We will also explore the microbial diversity in extreme terrestrial environments, as part of an effort to optimize the detection of non-abundant microbes. We will explore the microbes of desert, dry Antarctic valley, and stratosphere samples, which are models for high UV flux and desiccation. Michael Thomashow, an Astrobiology Institute PI at Michigan State University has offered to supply the Antarctic and Arctic samples.

The protocols and equipment will be tested on soil samples designed to mimic the Martian atmosphere and regolith, using UV fluxes and atmospheric gas pressures from Mars (41). Here the expertise of the Zuber lab will guide the selection of soils and radiation levels. These experiments are important because the reactivity of such oxidizing soils with the biochemicals in the PCR needs to be worked out. We must be assured that the Martian soil will not produce so many free radicals that the nucleotides or the DNA polymerase is inactivated. The use of free radical neutralizing chemicals will be explored.

Modular instrument design

There are multiple possible formats of the flight instrument, depending on how much weight and power is allotted. The minimal instrument would simply disrupt Martian soil, isolate DNA, and determine if a DNA segment of the proper length was amplified and after how many cycles. A slightly more complex instrument would pump any amplified DNA fragment to the another PCR station on the microfluidic chip, and query it with additional primers and another round of PCR. While the added specificity of this test is minimal, it also adds the ability to use nested PCR for additional sensitivity. The recursive PCR capability would add little weight to the microfluidic format. Finally the top-of-the-line instrument would include the gene array query/CCD reader module.

Engineering the other elements of the system

The disruption of soils by sonication should be simple to implement since it involves standard off the shelf components. The design of the soil receptacle and seal are significant but should not be a major hurdle once the robotic thermal cycler is developed. Similarly, we will address the heaters needed to liquefy the frozen water stores at -80 to 10 ° surface temperature.

Broadening the search if we do not detect a 16S RNA gene

a. sampling the unprocessed soil extracts

There are a number of points of redundancy for analysis, if no believable PCR signal is detected in the main DNA sample. First, the raw extract from the soil will be analyzed. For example, 0.1 microliters of the raw sample will be pumped to the PCR chamber for analysis. And increasingly large amounts of this extract could be added. Second, the particulate material in the 100 um to 1 um filters could be incubated for long periods of time in the protease/detergent extraction buffer, with multiple freezing during Martian nights, and thaws by heating the instrument area. Almost all microbes on Earth are about 1 um in diameter, so these filters would be expected to trap microbes if they escaped the sonication induced disruption. We will include valve connections between buffers and these filters to allow such a protocol if the main analysis sequence fails.

b. Using non 16S RNA gene oligonucleotide primers

Expanding the DNA sequence space searched by 16S primers

By carrying a few other DNA oligonucleotides other than universal 16S primers, we can broaden the search for DNA on Mars. For example, it is possible that life on Mars is very divergent so that one of the universal primers will work but not two of them. In such a case, PCR amplification using any <u>one</u> of the 4 primers, 505F, 906F, 906R, 1510R, and a set of xxx random primers yyy long **might yield a DNA duplex that would be detected with query DNA oligonucleotides, or the gene array. In this way, we would not be dependent on conservation of two regions of the 16S gene, but only one region. In a divergent 16S gene, this is much more likely.**

Searching for DNA with the random oligonucleotides.

It is also possible that there are microbes on Mars that use DNA but no longer bear a 16S gene detectable with the primers described above, either because they diverged before the invention of the 16S gene, deleted their 16 RNA gene, or have diverged in 16S gene sequence. In such a case, we would still be able to detect the amplification of DNA using a set of random short DNA oligonucleotides. The signature of life would be the detection of amplified DNA segments. The amplified DNA would be detected by SYBR Green fluorescence and its length could be roughly measured by a melting curve of the fluorescence or duplex. This protocol is much inferior to the 16S PCR because we cannot predict what DNA segment will be amplified nor what size will be amplified. Because of this, the source of the resulting DNA segment could not be determined unless we can optimize the 8mer DNA sequencing gene array. Without a DNA sequence, the only evidence that the resulting amplified DNA was of indigenous Martian origin would be its repeated isolation in Martian samples and scaling with the input amount of Mars sample. In every way this fall back analysis is inferior to the specificity of the 16S gene analysis, but it is included as an essentially no cost addition in case the 16S analysis fails. The only cost of its inclusion is the weight of the storage of a few ng of oligonucleotides and a few ug of the water to hydrate those oligonucleotides.