Analysis of Fungal Community Diversity in a Petroleum Contaminated Land Treatment Unit Using Terminal Restriction Fragment Patterns (TRFPs)

by

Kristyna E. Palmer

Biological Sciences Department

College of Science and Mathematics

California Polytechnic State University

San Luis Obispo, CA

December 2000

Table of Contents

List of Tables

List of Figures

Analysis of Fungal Community Diversity in a Petroleum Contaminated Land Treatment Unit Using Terminal Restriction Fragment Patterns (TRFPs)

Abstract

Terminal restriction fragment pattern (TRFP) analysis was used to assess the genetic diversity of fungi present in petroleum-contaminated soil. The soil was obtained from a pilot scale petroleum-degradation land treatment unit (LTU) studying the effects of growth conditions on degradation. Four separate cells, each treated with different growth conditions, were observed. This study will investigate the diversity of organisms in cell one, in which only nitrogen and phosphorus sources were supplied to the soil. To begin the analysis, PCR was carried out using DNA isolated from the soil samples and primers homologous to conserved regions in the fungal 18S ribosomal gene. The PCR products were digested with *Msp*I, *Hae*III, or *Dpn*II restriction enzymes. Because of the differences in the ribosomal DNA sequence of assorted organisms, the PCR reaction produced amplicons of assorted sizes with unique restriction cut sites. Therefore, digestion resulted in numerous labeled terminal DNA fragments (TRF) of various lengths. Capillary gel electrophoresis was used to analyze the lengths of the TRFs resulting in a distinctive terminal restriction fragment pattern (TRFP) for each sample. TRF patterns varied depending on the enzyme used in the restriction and the day the sample was collected. TRFPs were then compared to each other to observe shifts in diversity and fungal community members as a result of time, TPH levels, and conditions present. No known fungi were found that produced the exact same size TRFs as those found in our community. Overall, the TRFPs proved to be a useful tool in assessing the diversity of fungal species in the contaminated soil. With further development this technique shows promise in analyzing fungal and bacterial diversity in the environment which standard laboratory plating techniques cannot provide.

I. Introduction

Recently, interest in the study of biodegradation of pollutants in the contaminated environments by microorganisms has increased. By investigating the mechanisms these organisms possess that allow them to break down and use these toxic products as energy sources, more may be learned regarding the key to cleaning up disastrous accidents that lead to the destruction of our environment.

The contaminated site under study in this project is located in Guadalupe, CA. The site lies along the central coast of California extending into both San Barbara and San Luis Obispo counties. From 1946 to 1994 the site was used for the production of natural gas and oil. Unocal acquired total ownership of the site in 1953 and used the land until 1994 when all oil production ceased (5). During the period of oil production, diluent, which was used to aid in transport of the thick oil through the pipes, seeped into the surrounding environment. This diluent contained petroleum hydrocarbon in addition to other pollutants such as BTEX and PNAs (5). With funding provided by Unocal, the Environmental Biotechnology Institute at Cal Poly, San Luis Obispo used the unfortunate accident as an opportunity to oversee and study the bioremediation of the contaminated land. A pilot land treatment unit (LTU) was constructed in order to test the parameters and closely monitor and study the progress of the hydrocarbon degradation on a pilot scale. The pilot LTU was divided into four cells. Each cell was supplemented with sources of nitrogen and phosphorus from potassium phosphate, ammonium phosphate and ammonium sulfate. The soil was also monitored for moisture content and maintained between 6 and 10% water (5). In addition, the soil was tilled in order to supply adequate oxygen to the organisms (5). Since the reactions of biodegradation are mainly oxidation reactions, the process occurs

most efficiently if molecular oxygen is in abundant supply to the organisms (2). In cells 2, 3 and 4 a supplemental carbon source was added to the soil. Corn steep liquor was added to cell 2, dextrose was added to cell 3, and cornstarch was added to cell 4. Cell one was used as a control where no additional carbon source was added. These additions were made to see the effects of varying complexities of carbon sources to the degradation process.

Fungi and bacteria present in the soil are the organisms primarily responsible for the degradation of hydrocarbons in a contaminated environment (2). Complete bioremediation of a contaminated site requires a diverse collection of different organisms, including both bacteria and fungi. Each individual organism is only capable of metabolizing a specific range of hydrocarbon substrates (7). The microbes present in a contaminated environment, just as in any environment, will all complement each other's metabolic capabilities in order to take advantage of all of the possible energy sources. Because diversity is such a vital component of the dynamics of a particular environment, it is a critical factor to consider when researching the characteristics of a particular population.

The purpose of establishing the pilot land treatment unit was not only to investigate the differences of degradation between each of the cell conditions, but also to observe the numbers and diversity of organisms present in the contaminated environment and what, if any, shifts in these parameters occurred over time as conditions changed. However, accurately measuring the numbers and diversity of microorganisms in an environment is a difficult task. In the past, in order to observe the diversity of a community plating techniques were often used. Although it is a convenient and easy technique, it has been shown that plating methods fall short of desirable results and rarely show the true diversity of the organisms present (3). Plated soil samples will only provide data for organisms that can be cultivated using existing laboratory techniques (1).

Research has shown that observing stained samples of cells from the soil yields cell counts of one to two orders of magnitude larger than the plate counts (7). This indicates only about 1% of the actual organisms present in the sample are being accounted for in the measure of diversity using plates (7). In addition, a limited number of different colony morphologies for both bacteria and fungi result in difficulties in distinguishing between different species or even genus (1). Finally, any conditions that the organisms are subject to after collection may alter the community structure due to selection (3).

Several methods have been developed to attempt to more accurately access the diversity of complex communities of organisms. Many of these efforts rely on using the ribosomal genes that are present in both bacteria and fungi. These genes have been shown to contain divergent sequences between many species (4). Nearly all of these methods for community characterization include using PCR based protocols that amplify genes for analysis, followed by cloning and/or sequencing in order to identify differences in the DNA obtained from the sample (6). This process is often undesirable due to the lengthy process of cloning the thousands of amplicons into vectors, in addition to the time consuming amplification and sequencing (3). Another option for characterizing a community of organisms is to use a technique called amplified ribosomal DNA restriction analysis, or ARDRA (3). This method involves the restriction digest of amplified ribosomal DNA. The fragments are then visualized using a polyacrylamide gel to produce a community specific pattern (3). While this process is more rapid than cloning, it is limited by several problems. Since ARDRA uses visual techniques, small fragments or small amounts of particular fragments may be unresolved (1). These limitations led to the need for an alternate method that would be quick and inexpensive but also accurately characterize the population.

Another method, which also utilizes divergent sequences between different species in the ribosomal DNA to analyze community diversity, is terminal restriction fragment pattern (TRFP) analysis (1). This process uses a PCR reaction to amplify a region in the 16S ribosomal gene of bacteria using a fluorescently labeled primer and DNA isolated from the sample as the template. Amplification followed by restriction enzyme digestion are used to produce labeled fragments of various lengths. The combination of these fragments from the sample result in terminal restriction fragment patterns reflecting community diversity and abundance. These patterns have been shown to be useful in analyzing complex bacterial communities (1).

In this study, TRFP analysis was performed on soil samples obtained from Cell One of the pilot scale LTU. In order to determine if changes occurred in the total number and overall diversity of fungi in the community, primers homologous to regions in the fungal 18S ribosomal genes were used in the PCR reaction followed by restriction enzyme digests. The primers used in this protocol had been previously shown to amplify a 550 bp region that contains a zone of divergent sequence between different fungal species (8). TRFPs generated by this process are analyzed in this study.

.

II. Materials and Methods

DNA Extraction

Samples from cell one of the Guadulpe Land Treatment Unit were taken at days 14, 49 and 98. DNA was extracted separately from each sample using a modified Ultraclean MoBio Soil DNA Kit (MoBio Laboratories, Inc., Solano Beach, CA). The protocol for the use of the kit was followed to obtain 3 DNA samples, one for each collection day.

Gel Electrophoresis

A 1.5% agarose gel was prepared in TBE buffer, pH 8.5 (89.2 mM Tris, 88.9 mM boric acid, 2.47 mM disodium EDTA). Five _L of each sample and 1 _L of 5X loading buffer was loaded and the gel was run at 100v for 35 minutes. After staining with ethidium bromide, the gel was observed to check for successful DNA extraction from the soil.

DNA Quantification

The concentration of DNA in each of the three samples was measured using the SPECTRAmax UV spectrophotometer® (Perkin-Elmer, Applied Biosystems Inc., Fremont, CA). DNA concentrations are shown in Table 1.

Table 1. Calculations for PCR reaction

SAMPLE	[DNA]	DNA VOLUME (_L)	WATER (_L)
C1:14	57.24	17.5	82.5
C1:49	14.65	68.3	31.7
C1:98	5.57	100	0*

Need 10ng/_L Total final volume 100 _L

$C_1 V_1 = C_2 V_2$

*Note C1:98 is already very dilute so it will not be diluted further.

The concentration of the DNA in each sample was quantified by using SPECTRAmax UV spectrophotometer®.

Table 2. PCR Reagents and Volumes Required

Master Mix Reagent	**Volume (_L)**
Template DNA	6
10X Reaction Buffer	5.0
MgCl2, $1.75x10^{-4}$ mmoles	7.0
dNTPs, $3x10^{-5}$ mmoles	3.0
FUN18s1	1.1
FUN18s2	1.1
H2O	26.5
Taq DNA Polymerase, 2U	0.3
Total Reaction Volume	**50**

Calculations were done to find the volumes of each PCR reagent needed to provide the desired final concentration in the reaction.

PCR, 18S Labeled

PCR primers homologous to conserved regions of the fungal 18S ribosomal gene were designed to amplify a region of approximately 550 bp which is known to be divergent in sequence between species (7). The sequence of these primers are:

FUN 18S1 5_ CCATGCATGTCTAAGTWTAA 3_, forward

FUN 18S2 5_ GCTGGCACCAGACTTGCCCTCC3_, reverse.

DNA samples were diluted to obtain a final concentration of 10ng/_L. PCR reactions were carried out in triplicate using the reagents in Table 2.

The PCR was carried out using a thermal cycler. The reaction began with a 2 minute denaturation at 95°C. Then 40 cycles of the following series 95°C for 1 minute, 55°C for 1 minute, then 72°C for 2 minutes. The PCR was completed with a 10 minute cycle at 72°C to complete elongation.

Clean-up of PCR Products

A 1.5% agarose gel was run with 3.5 _L of each PCR reaction product and 1 _L of 5X loading buffer. After visualizing the gel, the successful PCR reactions for each sample were combined. The Ultra Clean-Up DNA Purification Kit (MoBio Laboratories, Inc., Solano Beach, CA) was used according to the protocol to remove any unused reactants and concentrate the amplicons.

Quantitation of Labeled PCR Product and Enzyme Digestion

The labeled PCR products were quantified using the SPECTRAmax UV spectrophotometer® as described above. The concentrations of the DNA in the PCR product are shown in Table 3.

Table 3. DNA Concentrations and Volumes Required for Enzyme Digest

SAMPLE	[DNA]	DNA VOLUME (_L)	WATER (_L)
C1:14	28.42	3.5	2.2
C1:49	23.138	4.3	1.4
C1:98	17.396	5.7	0.0

$C_1 V_1 = C_2 V_2$ C_1 =28.42, C_2 =10ng DNA/_L (desired volume), V_2=used volume of lowest [DNA]

The concentration of the DNA in each sample was quantified by using SPECTRAmax UV spectrophotometer®.

The PCR products were diluted with water in order to add 100 ng of DNA to the enzyme digest. The enzymes used in the digest were *Msp*I, *Dpn*II and *Hae*III. The restriction reactions had a total volume of 14.3 _L, and contained 100 ng of DNA and 2U of restriction enzyme. Three separate digests were performed on each sample. The samples were digested at 37ºC for 5 hours and then heated to 65ºC for 20 minutes to deactivate the enzymes.

Ethanol Precipitation of Digest

To each digest, 50 _L of cold 95% ethanol and 1 _L 3M sodium acetate (pH 4.6) were added and mixed. The mixture was incubated at 4ºC for 30 minutes and then centrifuged for 15 minutes at 15,000 to pellet the DNA. The ethanol was removed and the pellet was washed with 50 _L of cold 70% ethanol, centrifuging again for 15 minutes at 15,000. The ethanol was removed and the pellet of DNA was dried by vacuum centrifugation.

Capillary Electrophoresis

The DNA pellet was resuspended in 20 _L of formamide (BioRad, Benecia, CA), 1 _L of load buffer and 1 _L each of Genescan 500 ROX® (Perkin-Elmer, Applied Biosystems Inc., Fremont, CA) and ROX® 550-700 (BioVenture, Murfreesboro, TN) standards. The samples were denatured at 95ºC for 5 minutes and placed directly into an ice bath for 10 minutes to snap cool. This insures that the DNA remains denatured. The samples were then run on an ABI PrismTM 310 Genetic Analyzer (Applied Biosystems Inc., Fremont, CA) at 15 kV at 60ºC. Genescan 3.1 software (Applied Biosystems Inc., Fremont, CA) was used to analyze the output from the Genetic Analyzer. Using the size standard peaks, the electropherogram peaks were resolved to one base pair. The result is that the labeled terminal restriction fragments present in each sample can be seen on the electropherogram as an individual peak, where the height of the peak represents the relative amount of that fragment in the digested PCR product.

Data Analysis

Data truncation and normalization were used to compare the three samples to each other. Since the amount of DNA loaded onto the ABI Prism 310 can not be accurately controlled, area varied between TRFPs for the different samples. In order to justly compare the different TRFPs to each other, the peak detection threshold had to be truncated thereby creating new peak detection thresholds for each sample. Truncation was accomplished by first calculating the ratio of each peak area to the area of the smallest peak and then multiplying that ratio by the smallest peak area the ABI Prism 310 can detect (580 area units). The peaks that had areas lower than this new threshold level were then removed from the data. Table 4 illustrates how the truncation of the data was preformed.

After truncation the remaining peaks had to be normalized. To do this each peak area was converted to parts per million (ppm) by calculating the ratio of each peak area to the total area and multiplying by $1x10^6$.

Table 4. Example of Truncation Used to Prepare TRFP Data for Analysis

Sample	**Total Area**	**Ratio of Total Peak Area****	**Peak Area Analyzed*****
Smallest Peak	200000	1:1	580*
Peak	1000000	5:1	2900
Peak	2000000	10:1	5800

* Minimum detectable peak area with ABI Prism™ 310
** Ratio = Peak Area / Area of Smallest Peak
*** Peak Area Analyzed = ratio * 580

After completing the amplification and digestion, more DNA was present in some samples than in other samples. This resulted in TRFPs that had peaks much larger than the peaks generated from samples with less total DNA. In this investigation, the total amount of DNA was not important but rather the ratios of the different size DNA fragments. For this reason the peaks were all truncated in order to accurately compare peaks from different TRFPs.

III. Results and Discussion

After producing TRFPs for each of the three samples for each of the three enzymes, the patterns were compared to help characterize the population over time. Observing the peaks and clusters of peaks of the TRPFs from each of the three days provided an idea of how the diversity changed over time (see Figures 1, 2, and 3). Peaks occurring at or near 500 base pairs generally indicated no restriction of the fragment by the enzyme, either because of incomplete digestion, or due to a lack of restriction sites in the fragment. Therefore, peaks that are 500 bp or greater were not be analyzed.

Figure 1 shows the TRFPs generated by the fragments produced after restriction with *Msp*I. First, in the TRFP for day 14, the most significant peak is at 185 bp. There are also some smaller peaks centered at 143, 212, 324, 335, and 425 bp. At 14 days, the organism(s) that result in the 185 bp fragment are dominant in the population and those that result in the smaller peaks are more inferior in the community. By looking at the TRFP for the day 49 sample, it is apparent that the dominant species in community has shifted to the organism(s) which produce the peak at 233 bp. This peak was completely absent from the first sample at 14 days. This may indicate that these fungi were only able to grow or were only present after day 14 when certain conditions in the environment changed such as the level of hydrocarbon. The earlier dominant organism, which produced the 185 bp peak, completely disappeared from the community by day 49, as have many of the smaller peaks. At day 49, there is another large peak at 402 bp that was not present at day 14. The other minor peaks in this TRFP include those at 143, 213, 237, and 331 bp. The peak at 143 bp was present in both samples from days 14 and 49 and is present in almost the exact same quantity in both samples. This same trend is visible in the 212/213 bp

peaks of the day 14 and day 49 TRFPs. This indicates that this fungus or group of fungi may be a consistent, stable member of this community. Also, the fungi that produced the large peak in the day 49 sample at 233 bp is still present in the day 98 sample but in less quantity. This may be attributed to the fact that the energy source for this species is in short supply. Finally, in the day 98 sample there are several equally dominant fungi including those producing fragments at 173, 185, 324 and 425 bp. However, the largest peak for this sample is at 185 bp. This peak, while absent from the day 49 sample, was also the dominant peak in the day 14 sample. Peaks at 324 and 425 bp that were present in the day 14 TRFP, also follow this pattern of disappearing in the second TRFP and reappearing in the third.

Analysis of the TRFPs generated by *Dpn*II digest (Figure 2) shows that at day 14 three dominant peaks are present at 213, 218, and 476 bp. Of these three large peaks, only the one at 218 bp is still apparent at 49 days, although it is much reduced. Also, at day 49 there is a new large peak at 69 bp. However, the organisms that result in this peak are completely gone by day 98. The peak at 213 bp, which was present in the first TRFP, was not seen on day 49 but reappeared in the last TRFP. Again, the TRFP for day 98, also has a greater number of significant peaks as did the *Msp*I day 98 TRFP, which may indicate the increase in diverse organisms as the hydrocarbons are degraded and dead organisms accumulate.

The TRFPs generated by using the *Hae*III enzyme, shown in Figure 3, have numerous peaks that are spread over the fragment size region. The TRFP for day 14 show several large peaks at 135, 221, and 282 bp. The peak at 135 bp is still present at day 49, but is much smaller. Also a small peak at 143 bp is present on both TRFPs for days 14 and 49 and is of similar size in both. The organisms that produced the fragments of sizes 225 and 288 bp dominate the sample at day 49, however at 14 days the 225 bp peak was only barely detectable and the peak at 288 bp

was nonexistent. Finally, at day 98 the peak at 225 bp is smaller but is still one of the several dominant peaks of the TRFP and 288 bp has become a much less dominant member of the community. The new dominant peaks present in this TRFP in addition to those at 225 and 288 bp are the peaks at 257 and 281 bp.

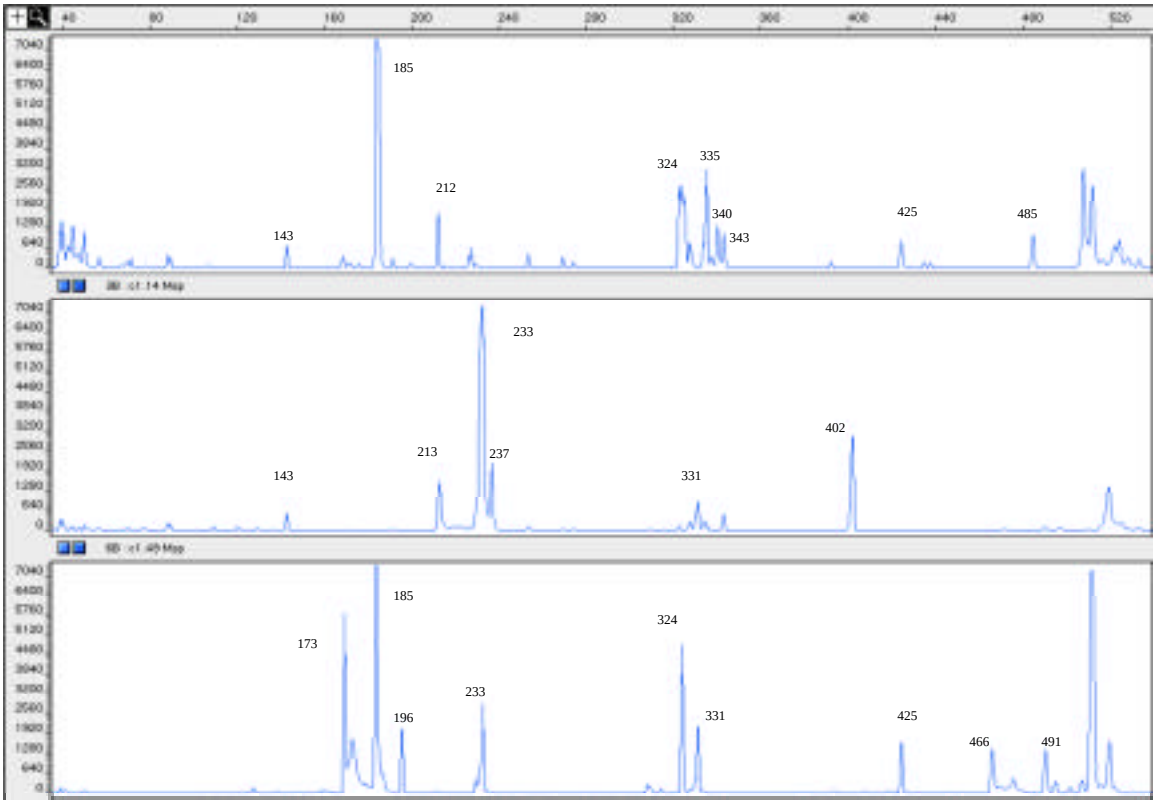
The incidence of disappearing and reappearing peaks my be due to the fact on day 14 certain organisms have an ample supply of their required energy sources, such as specific size hydrocarbon present in the oil. But as these fungi use up this substrate their numbers decline indicated by the lack of a peak at this size on day 49. However, if conditions change due to the tilling of the soil or metabolism of other organisms, the availability of the substrate used by this organism may increase again resulting in the reappearance of this fungus. In all of the day 98 samples there are a greater number of significant peaks in this sample, this suggests a greater fungal diversity. This may be attributed to the fact that at this point in time the levels of total petroleum hydrocarbon in the soil had fallen to a level that more organisms are able to tolerate (see Figure 5). In addition, as bacterial levels in the soil drop there is a large supply of dead bacteria on which the fungi can grow.

Figure 1. TRFP of Cell 1 Samples from Days 14, 49, and 98 Digested with *Msp*I

Day 14

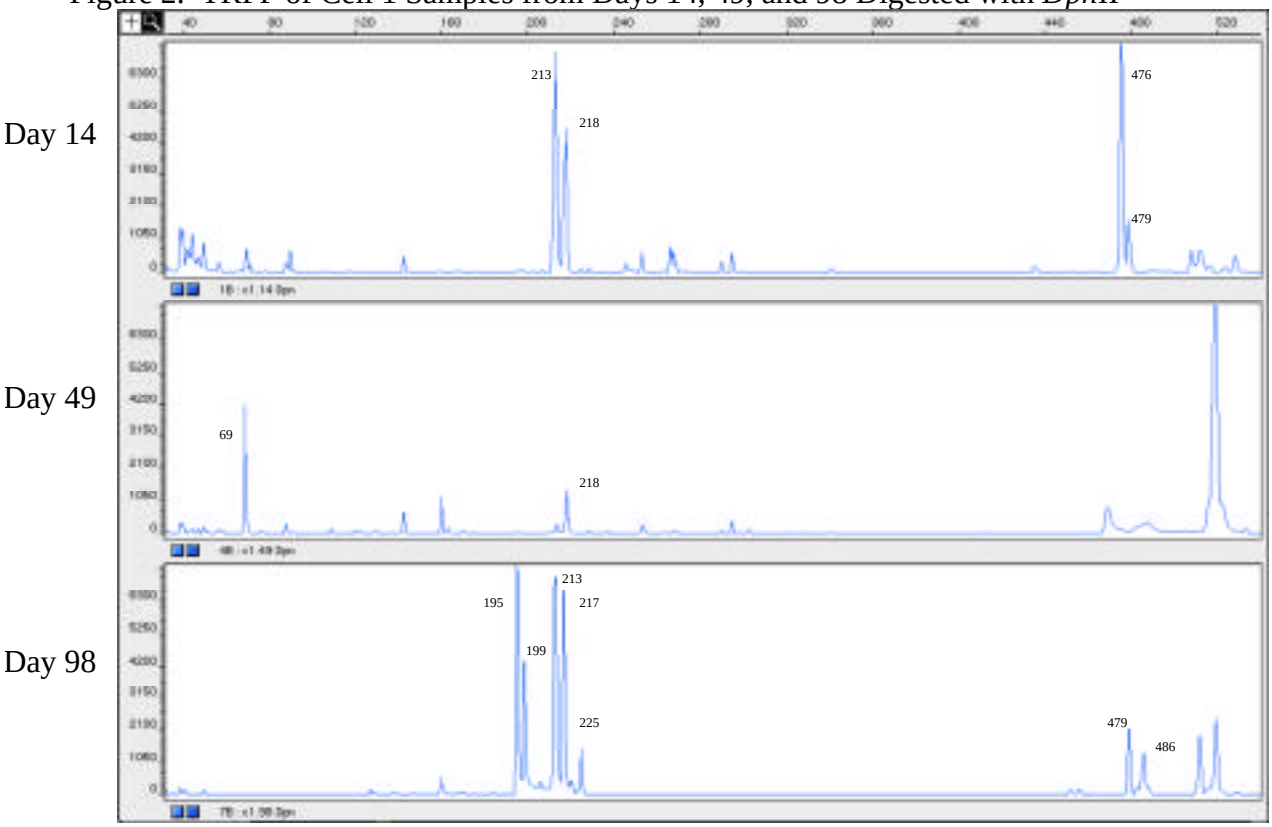
Day 49

Day 98



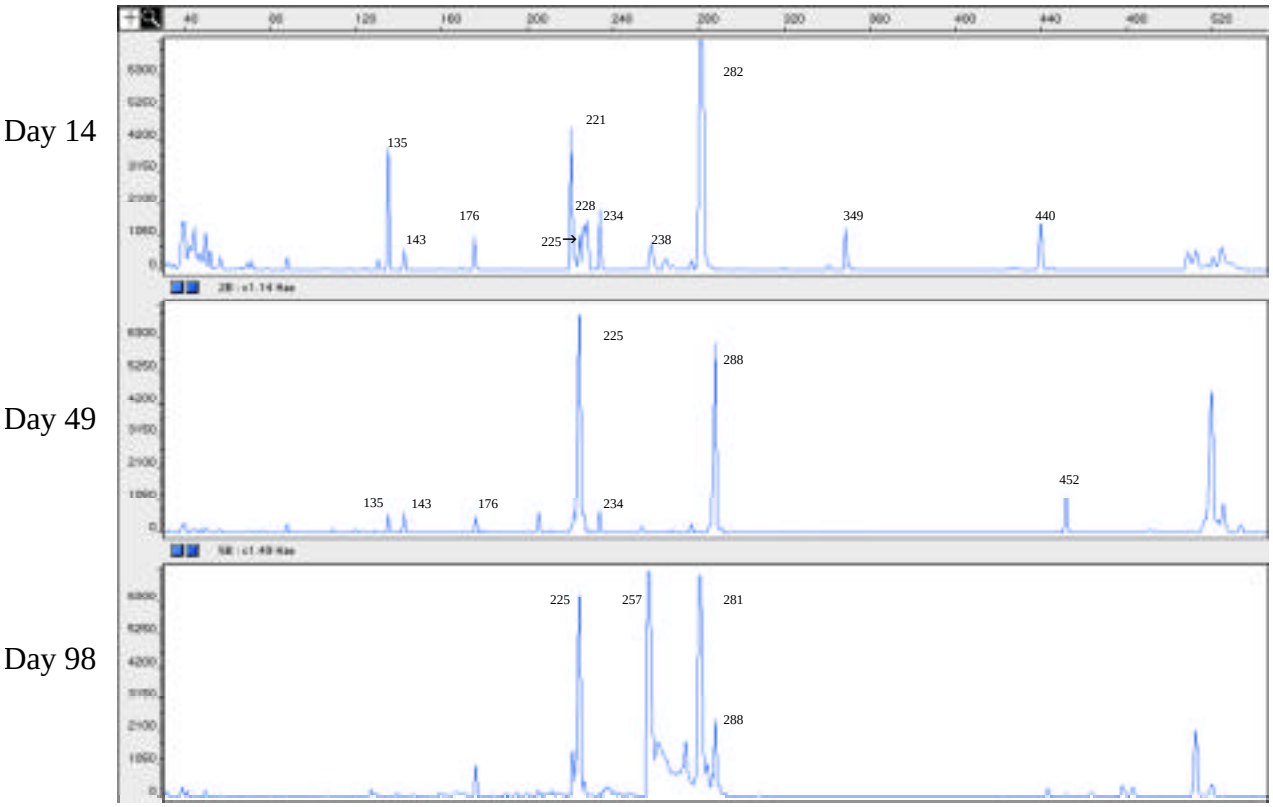
This electropherogram was generated by applying the Genescan 3.1 software to the data generated by the capillary gel eletrophoresis. Each peak represents one of the lengths of TRFs a generated by digestion with *Msp*I. The height of each peak represents the relative amount of each TRF.

Figure 2. TRFP of Cell 1 Samples from Days 14, 49, and 98 Digested with *Dpn*II



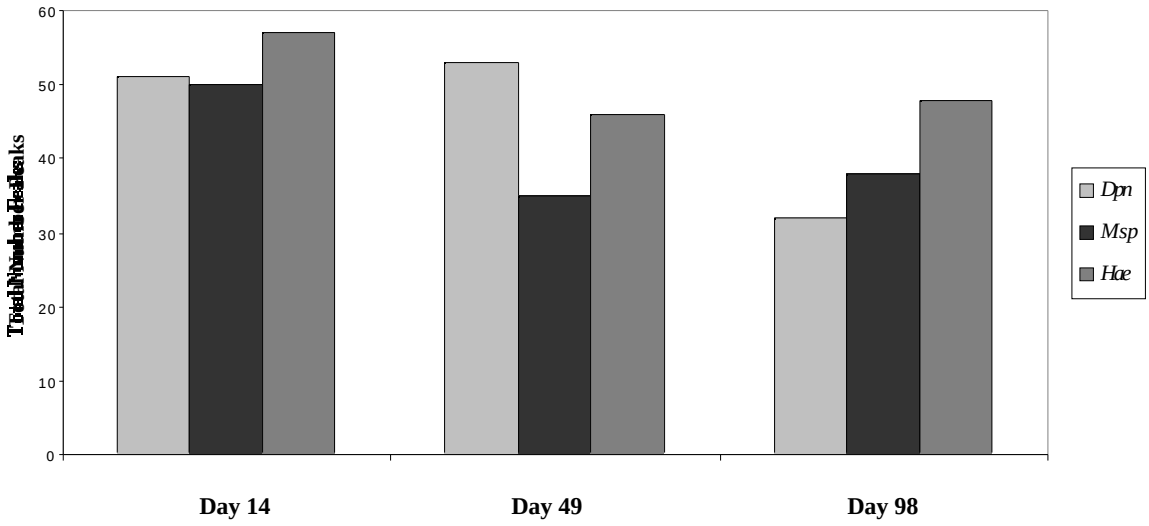
This electropherogram was generated by applying the Genescan 3.1 software to the data generated by the capillary gel eletrophoresis. Each peak represents one of the lengths of TRFs a generated by digestion with *Dpn*II. The height of each peak represents the relative amount of each TRF.

Figure 3. TRFP of Cell 1 Samples from Days 14, 49, and 98 Digested with *Hae*III



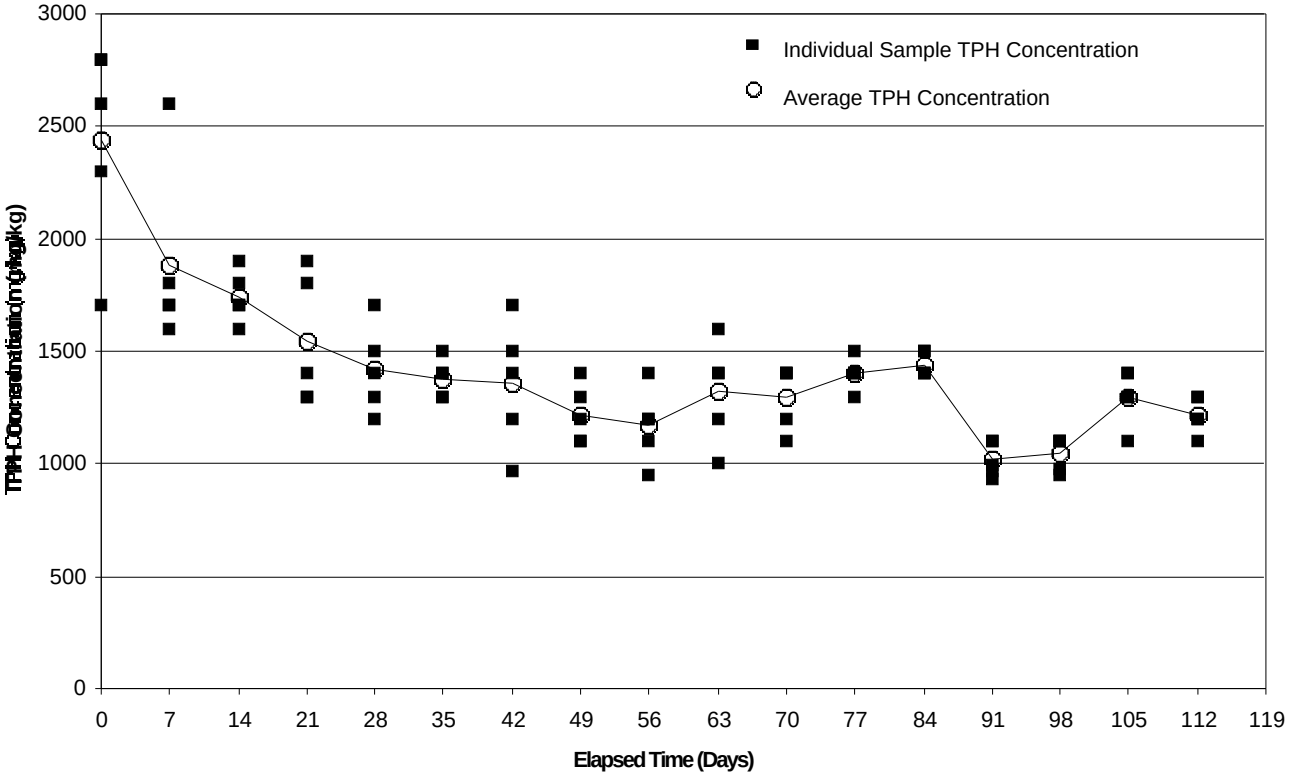
This electropherogram was generated by applying the Genescan 3.1 software to the data generated by the capillary gel eletrophoresis. Each peak represents one of the lengths of TRFs a generated by digestion with *Hae*III. The height of each peak represents the relative amount of each TRF.

Figure 4. Total Peaks Detected in TRFPs for each enzyme and in each sample



This graph was generated using the total number of peaks of all sizes and quantities generated by the three different enzyme digests. These numbers are a representation of total diversity in the community.

Figure 5. Total Petroleum Hydrocarbon Averages for Cell 1



This graph was generated using data obtained during the pilot study of petroleum hydrocarbon degradation in cell one. The graph was acquired from the Pilot Land Treatment Unit Report (5). These data points represent the total level of all hydrocarbons with various carbon chains lengths present in the samples.

Figure 4 shows the relative number of peaks per TRFP for each day and each enzyme. These numbers reflect every peak present including the smallest peaks. This graph gives an idea of how the diversity levels of the community changed over time. For both *Msp*I and *Hae*III the number of peaks both follow a pattern of a decrease during the middle sample and then an increase again for the final sample. This pattern is opposite in the case of the *Dpn*II generated TRFPs. The changes in diversity that were monitored using this new technique seem to suggest that diversity of fungi in the contaminated community is constantly changing depending on the environmental conditions. In this contaminated cell, the fungi were constantly shifting in order to account for the changing environmental conditions. In the case of hydrocarbon degradation, some organisms are able to metabolize long chain hydrocarbons while others utilize short chains. Other organisms are also present which utilize organic material from the other dead cellular matter. All of these factors and others cause different organisms to thrive at different points of the bioremediation process, and TRFPs allow the monitoring of these changes in diversity.

Tables 5, 6 and 7 show various organisms that may be present in the soil samples for each of the three days. A program was run on the peak data for all the TRFPs from each day comparing them to expected peaks sizes of specific organisms using databases containing sequences for the fungal ribosomal DNA. The fragment sizes formed in the TRPF process for each enzyme are then compared to these predicted fragment sizes. An organism that has predicted peaks for all three of the enzymes that match peaks present on the TRFPs for a particular sample is a likely match for an unknown organism present in that sample. The fungi listed in the tables below have predicted peaks that are similar to those present in the TRFPs for the sample within at least 20% of the size of the fragments. A deviation of only 10% would indicate a very strong likelihood that an organism is present in a sample. The predicted peaks for

C. apollinis in the Day 49 sample only deviated by 10% from the peaks in the TRFPs for that sample (see Table 3).

It is apparent that while these comparisons do not result in exact matches, fungal identities can be suggested. However, it cannot be forgotten that databases on fragment sizes do not exist for all organisms and that many of the fungi actually present in the sample may not have been cultured before. Also, an additional problem exists in the fact that often numerous organisms produce fragments of the same size.

By looking at the organisms in these three tables, the only genus that appears in multiple days is *Bensingtonia*, which appears in both the day 14 and day 49 tables. For day 14, this genus has predicted peak sizes of 135, 324, and 482 bp for *Hae*III, *Msp*I, *Dpn*II respectively. The peak at 135 bp for *Hae*III is visible on Figure 3, and the 324 bp peak for *Msp*I is shown on Figure 1. However, the predicted peak for *Dpn*II at 482 bp is not present in this TRFP. Although, by looking at Figure 2, there is a peak at 479 bp. Two of the TRFPs for day 14 contained peaks that exactly matched predicted fragment sizes for *Bensingtonia*. However, because the TRFP for *Dpn*II had a peak that was very close to, but not exactly, the same size as the third predicted fragment size, this organism was given a 20% deviation rating. This same process can be used to compare the predicted peaks sizes of other organisms in the tables to the peaks present in the community TRFPs.

Table 5. Organisms with predicted peaks within 20%* of those in the TRFs for Day 14

Organism	***Hae*III**	***Msp*I**	***Dpn*II**
Bensingtonia subrosea (20%)	135	324	482
Bracteacoccus (20%)	79	235	173
Carteria crucifera (20%)	170	234	490
Chlorella fusca (20%)	170	234	490
Endosymbiont (20%)	276	320	205
Lentinus tigrinus (20%)	521	144	521
Nitzschia pungens (20%)	276	320	205
Scenedesmus obliquus (20%)	170	234	490
Tetraselmis striata (20%)	169	234	491

This table was generated using a macro (Environmental Biotechnology Institute) that compared the data from the day 14 TRFPs to a database of predicted fragment sizes for particular organisms digested by the three enzymes. The result is a list of organisms that have predicted peak sizes for each of the three enzymes that are similar to the peaks seen in the TRFPs. The percentages shown indicate the deviation of the predicted peaks from the peaks in the TRFPs.

Table 6. Organisms with predicted peaks within 20%* of those in the TRFs for Day 49

Organism	***Hae*III**	***Msp*I**	***Dpn*II**
Auxarthron zuffianum (20%)	222	230	517
Bensingtonia musae (20%)	176	224	*481*
C.apollinis (10%)	221	229	515
Candida krusei (20%)	179	225	*479*
Capronia mansonii (20%)	*222*	230	517
Chrysosporium parvum (20%)	222	230	517
Cordycepioideus bisporus (20%)	222	230	517
Cordyceps intermedia (20%)	222	230	517
Cryphonectria cubensis (20%)	222	230	517
Cryphonectria havanensis (20%)	222	230	517
Cryphonectria parasitica (20%)	222	230	517
Cryphonectria radicalis (20%)	222	230	517
Ctenomyces serratus (20%)	222	230	517
Disciotis venosa (20%)	222	230	517
Endothia gyrosa (20%)	222	230	517
Exophiala dermatitidis (20%)	222	230	516
Exophiala dermatitidis (20%)	222	230	518
Exophiala mansonii (20%)	222	230	517
Gymnoascoideus petalosporus (20%)	222	230	517
Hypocrea lutea (20%)	222	230	517
Hypomyces chrysospermus (20%)	222	230	517
Leucostoma persooni (20%)	222	230	517
Malbranchea dendritica (20%)	222	230	517
Microbotryum violaceum (20%)	222	230	161
Obertrumia georgiana (20%)	515	327	515
Paecilomyces tenuipes (20%)	222	230	517
Protomyces inouyei (20%)	222	230	516
Protomyces lactucae (20%)	222	230	516
Renispora flavissima (20%)	222	230	517
Sarcinomyces phaeomuriformis (20%)	222	230	518
Sclerotinia sclerotiorum (20%)	222	230	517
Taphrina deformans (20%)	*222*	230	517
Taphrina populina (20%)	222	230	516
Tilletiopsis washingtonensis (20%)	222	230	517
Verpa bohemica (20%)	222	230	517
Verpa conica (20%)	222	230	517
Wangiella dermatitidis (20%)	*222*	230	517

This table was generated using a macro (EBI) that compared the data from the day 49 TRFPs to a database of predicted fragment sizes for particular organisms digested by the three enzymes. The result is a list of organisms that have predicted peak sizes for each of the three enzymes that are similar to the peaks seen in the TRFPs. The percentages shown indicate the deviation of the predicted peaks from the peaks in the TRFPs.

Table 7. Organisms with predicted peaks within 20%* of those in the TRFs for Day 98

Organism	***Hae*III**	***Msp*I**	***Dpn*II**
Leucosporidium scottii (20%)	177	330	161
Mallomonas adamas (20%)	281	186	215
Mallomonas akrokomos (20%)	282	186	215
Mallomonas annulata (20%)	282	186	215
Mallomonas splendens (20%)	282	186	215
Mallomonas striata (20%)	281	186	215
Paraphysomonas foraminifera (20%)	284	188	217
Rhodotorula glutinis (20%)	177	330	161
Spizellomyces acuminatus (20%)	227	178	522
Symbiodinium (20%)	229	329	515
Synura spinosa (20%)	282	186	215
Synura uvella (20%)	282	186	215
Tessellaria volvocina (20%)	282	186	215

This table was generated using a macro (Environmental Biotechnology Institute) that compared the data from the day 98 TRFPs to a database of predicted fragment sizes for particular organisms digested by the three enzymes. The result is a list of organisms that have predicted peak sizes for each of the three enzymes that are similar to the peaks seen in the TRFPs. The percentages shown indicate the deviation of the predicted peaks from the peaks in the TRFPs.

IV. Conclusion

By observing the results obtained in this study, it is apparent that restricting the fragments with different enzymes can yield very different TRFPs. By comparing the total number of peaks in the different TRFPs to each other, the overall changes in the level of diversity in the community could be estimated. Figure 4 depicts the total number of peaks for each of the TRFPs. This graph shows that although *Msp*I and *Hae*III indicate a decrease and then an increase in the community diversity over time, the *Dpn*II data indicates the opposite. The reason that *Dpn*II did not result in the same pattern for diversity shifts as the other two enzymes is unknown and may be a result of a significant difference in restriction site patterns. Assuming that the pattern represented by both *Msp*I and *Hae*III is correct and the diversity decreased, and then increased again, some speculation can be made about what took place in the community during the period of the study. In the beginning (day14), there was an abundance of nutrients and energy sources available to the organisms so the diversity was high, but as these substrates were used up and conditions became less favorable, many organisms died off, leaving only those few fungi that were able to utilize the most complex hydrocarbons to survive (day 49). After these fungi had broken the complex hydrocarbons into more simple ones and the total petroleum hydrocarbon levels had dropped, more fungi could tolerate the conditions and the diversity of the community increased again (day 98).

While analysis of TRFPs generated using the 18S ribosomal DNA showed changes in diversity of the community, this technique was limited in predicting what organisms were present in the sample. After producing TRFPs for several isolated cultures, it was discovered that many fungi produced TRF peaks of the same size. This can also be observed in the predicted peak sizes for various organisms in the database. Tables 5, 6 and 7 all show that there

are many organisms that produce the same size TRFs using the same enzyme. In other cases, it was discovered that multiple peaks could also be present for a single organism. This occurrence was caused by an incomplete digestion of fragments with multiple restriction sites. Another problem with the TRFPs was that in some cases the peaks were very close together resulting in difficulty distinguishing the separate peaks, and therefore the separate organisms, from each other. Restriction with *Hae*III seemed to produce a greater variety of peak sizes than the other two enzymes. With further investigation, an enzyme that results in better peak separation could be discovered.

Besides using different enzymes, another possible solution to these problems is to use a region which is less conserved between species, such as the ITS region, for forming the TRFs. This region may contain more polymorphisms in the sequence between different species and therefore may produce a greater variety of fragment sizes. By finding the most effective enzyme and a less conserved region of the fungal genome, TRFPs could become very useful for identifying the presence of specific organisms in a community.

Although this technique has not yet been perfected so that it can be used to accurately identify the organisms present in the sample, it can be used to make educated guesses. The fungi that showed predicted peaks close to those that appeared in the TRFPs from each sample are in Tables 2, 3, and 4. However, none of the organisms match exactly to the peak sizes present in the TRFPs. This may be due to the fact that the organisms present in the sample are either not in the database, or have never been cultured. In addition, the fragments may have lost some base pairs due to physical damage during the PCR reaction. With further research this technique may eventually be able to describe the exact organisms present in a community.

Terminal restriction fragment pattern analysis is the first step to obtaining a more complete understanding of community dynamics. This technique provides a way to observe and monitor the diversity of a community in which all members of that community are included in the analysis. TRFP analysis also allows researchers to study communities in which the organisms are difficult to culture, such as extremely fastidious microbes, or those that require specific growing conditions to survive culturing, such as high temperatures. Continued study of the diversity of fungal communities in soil will be useful in many applications, in addition to the future of environmental bioremediation. Knowing the optimal conditions necessary to obtain the greatest diversity and highest level of hydrocarbon degradation may help to maximize petroleum degradation. By achieving a high level of diversity, a greater variety of the toxic contaminants may be degraded. Although there are strict measures in place to prevent accidents that result in environmental contamination, when these accidents do occur research in bioremediation will help to restore the environment to its natural state.

Acknowledgements

This project would not have been possible with out the help of Dr. Susan Elrod, the project advisor, and Chris Kaplan of the Environmental Biotechnology Institute. The work for this project was financially supported by grants from Unocal.

References

2. Clement, Brian G., Lucia E. Kehl, Kristin L. DeBord, Christopher L. Kitts. *Terminal restriction fragment patterns (TRFPs), a rapid PCR-based method for the comparison of complex bacterial communities*. Journal of Microbiological Methods. 31 (1998) 135-142.
3. Leahy, Joseph G., Rita R. Colwell. *Microbial degradation of hydrocarbons in the environment*, Microbiological Reviews. 54 (1990) 305-315.
4. Liu, Wen-Tso, Terence L. Marsh, Hans Cheng, Larry J. Forney. *Characterization of microbial diversity by determining terminal restriction fragment polymorphisms of genes encoding 16S rRNA*. Applied and Environmental Microbiology. 63 (1997) 4516-4522.
5. Moffet, Walsh. *Analysis of bacterial community structure using 16S rDNA analysis*. Anaerobe. 6 (2000) 129-131.
6. Pilot Land Treatment Unit Report Guadalupe Oil Field, San Luis Obispo, CA. February 25, 2000. Environmental Biotechnology Institute, California Polytechnic State University, San Luis Obispo, CA.
7. Rondon, Michelle R., August, P.R., Betterman, A.D., Brady, S.F., Grossman, T.H., Liles, M.R., Loiacono, K.A., Lynch, B.A., MacNeil, I.A., Minor, C., Tiong, C.L., Gilman, M., Osburne, M.S., Clardy, J., Handelsman, J., Goodman, R.M. *Cloning the soil metagenome: a strategy for accessing the genetic diversity of uncultured microorganisms*. Applied and Environmental Microbiology. 66 (2000) 2541-2547.
8. Rondon, Michelle R., Robert M. Goodman, Jo Handelsman. *The Earth's bounty; assessing and accessing soil microbial diversity*, Trends in Biotechnology. 17 (1999) 403-409.
9. Shank, Pam. *Fungal community analysis by terminal restriction fragment pattern for oil contaminated sand*. Senior Project, California Polytechnic State University, San Luis Obispo, Biological Sciences Department (1999).