THE EFFECT OF WILLOW TREE GROWTH ON BACTERIA IN HYDROCARBON CONTAMINATED SOIL: A POSSIBLE ROLE IN PHYTOREMEDIATION.

By Richard Phan

Biological Sciences Department
College of Science and Mathematics

California Polytechnic State University
San Luis Obispo

Abstract

Phytoremediation is a process involving the use of green plants to safely and naturally degrade harmful chemical pollutants that have been introduced into an environment. In this study, samples were collected from a volunteer willow tree found growing in a mound of soil that was heavily contaminated with petroleum. The goal of this study was to characterize and compare the composition of the bacterial community in the soil surrounding the willow tree to samples elsewhere in the stockpile. Three sets of samples were obtained on April 2004 August 2004 and December 2004 at 2, 4, and 8 ft depths. Aerobic heterotrophic bacteria were estimated by direct plate count on R2A agar. The microbial community was analyzed by observing Terminal Restriction Fragment (TRF) patterns of the bacterial 16S ribosomal genes. DNA was extracted from all samples and amplified by the Polymerase Chain Reaction (PCR) using primers homologous to the 16S ribosomal RNA genes of Bacteria. A restriction endonuclease digestion of the amplicons was carried out using one of two digestion enzymes *Dpn*II or *Hae*III. The digested fragments were separated by capillary gel electrophoresis to produce TRF patterns that were then analyzed. At first the microbial community structure in the willow soil was similar to the controls. However, the difference between the willow and control soils gradually increased over time. In addition, the aerobic heterotrophic bacterial count increased in the willow soil compared to the control soil and petroleum levels significantly dropped in the willow soil. We can conclude that the presence of the willow in the petroleum-affected soil is changing the microbial community in the soil, which is having an effect on the petroleum levels in the soil. A prolonged study of this willow tree could provide more specific information on how the willow tree is working to stimulate the microbial community, which in turn is lowering the TPH levels in the stockpiled soil.

Introduction

The inadvertent discharge of petroleum during oil production and shipping, greatly contribute to the environmental contamination of both marine and land environments. According to the EPA, there are as many as 450,000- 600,000 petroleum contaminated "brownfield" sites in the continental United States alone. These "brownfield" sites are mainly composed of abandoned industrial sites with one or more leaking underground petroleum storage tanks, or USTs (1). The Guadalupe dune restoration project, although not classified as a "brownfield" site, sustained massive hydrocarbon releases due to long term leaking of storage tanks containing petroleum hydrocarbons.

Bioremediation is the use of plants or bacteria to naturally break down harmful substances into less harmful byproducts. For example, some soil bacteria use petroleum hydrocarbons as a source of carbon, and ultimately produce cellular energy, with water and carbon dioxide as waste products. In contrast, the use of plants, or phytoremediation, may use a plant's ability to absorb the harmful chemicals. In addition, phytoremediation may also work to stimulate bacterial communities that facilitate the degradation of hydrocarbons, an added benefit compared to the use of bacteria alone.

From 1951 to 1994 the Guadalupe was an active oil field under the management of Union Oil Company of California (UNOCAL). However, the oil at the site was found to be very viscous, so "to enhance the flow characteristics of the crude, two main methods were used; diluent mixing and steam injection." (2). Diluent is a petroleum distillate what was used to thin out the oil and make it easier to pump, "Diluent is similar to kerosene/diesel mixture and contains low levels of volatile compounds (e.g. benzene, toluene, ethyl benzene, and total xylenes [BTEX] that are frequently associated with petroleum products." (2). Throughout the period in which the diluent was used, 1955-1990, there were leaks throughout the oilfield, resulting in large areas of affected groundwater and soil. Several of the affected areas were excavated to immediately limit environmental degradation and the affected soil was stockpiled while remediation options were investigated.

In 2003, a volunteer willow tree was discovered to be growing on the stockpile of excavated diluent affected soil. In this study, soil samples were taken and analyzed to determine the effect the willow tree had on both the microbial community, and the total petroleum hydrocarbon (TPH) levels in the soil. The goal of this study was to first characterize the

microbial community in soil samples near the willow and soils farther away from the willow. The willow was monitored over an 8-month period with samples analyzed in 4-month intervals. TPH levels were quantitatively monitored, aerobic heterotrophic bacteria counted, and the microbial community was characterized using Terminal Restriction Fragment length polymorphism analysis.

Materials and Methods

Sampling

Samples that were considered to be within the effect of the willow tree were taken 4 feet from the base of the tree and regarded as the willow sample. Samples that were considered to be out of the range of the willow tree were taken 20 feet away from the base of the tree and regarded as the control sample. The soil collected from these two sample sites was taken at three depths 2 feet, 4 feet, and 8 feet. Each sample was processed and analyzed for various chemical and physical characters including TPH. Small soil vapor extraction wells were placed near sampling sites to measure soil oxygen, methane and carbon dioxide levels.

Cell count by direct plating method

Samples were plated R2A agar (Beckton Dickenson, New Jersey) with serial dilutions of 10^{-5}, 10^{-4}, 10^{-3}, and 10^{-2} made using Bushnell-Haas media. Ten microliter aliquots were streaked in a straight line and incubated at 28°C for a 2-week period. Plates were counted in 1-week intervals, colonies counts were averaged for the 2-week period and CFU/g determined.

DNA extraction

DNA extraction was preformed on approximately 1 g of soil from each sample following the procedures provided by the MoBio Powersoil DNA kit. DNA was confirmed by gel electrophoresis on a 1.5% agarose gel run with TBE buffer.(82.9mMtris, 88.9mM boric acid, 2.47mM disodium EDTA). The gel electrophoresis ran at approximately 106V for approximately 30 minutes and then immersed in ethidium bromide for approximately 30 minutes. The gel was visualized using the Bio-Rad Gel Doc system.

PCR amplification

PCR was carried out using primers specific for the highly conserved eubacterial 16S ribosomal genes. The primers used in this study were Ba2F, a forward primer that was flourescently labeled with phosphamide dye and K2R, the reverse primer (Both provided by Applied Biosystems, Fremont, CA). The PCR reactions consisted of 1μL of extracted DNA containing approximately 1 ng/μL, 29.7 μL of purified deionized water, 5 μL of 10X Buffer, 3 μL of 10 mM

DNTP's, 2 µL of 20 µg/ml BSA, 7 µL of 25 mM $MgCl_2$, 1 µL labeled Ba2F, 1 µL K2R, and 0.3 µL of 5 U/µL *Taq* gold enzyme. All reactions were placed in a thermocycler and heated to 94 °C for 10 minutes, followed by 30 cycles of 94°C for 1 minute, 46.5°C for 1 minute, and 72°C for 2 minutes. Each sample was run in triplicate. Once completed, the samples were held at 72°C for 10 minutes and held at 4°C until samples were ready to be used.

PCR clean up

To confirm the presence of the 16S amplicons, gel electrophoresis was run as above. Confirmed samples were then concentrated and purified using the Mo Bio Ultraclean PCR cleanup kit following the protocols accompanied by the kit. Samples were then quantified using a UV Flourometer.

Restriction Endonuclease Digestion

After PCR clean up, the samples were digested using one of two restriction endonuclease enzymes provided by New England Bio labs, *DpnII* or *HaeIII*. Each sample was digested using only one enzyme but both enzymes were using for every sample in a set. For the enzyme *DpnII* the reaction consisted of approximately 70 ng of amplified DNA mixed with 4 µL of *Dpn* buffer, and 0.5 µL of the *Dpn* enzyme. For the enzyme *HaeIII*, the reaction consisted of approximately 70 ng of amplified DNA mixed with 4 µL of Buffer 2, also provided by New England Biolabs, and 0.2 µL of the enzyme. Each Digestion was incubated at 37°C for 4 hours and then enzyme was deactivated at 65°C for 20 minutes for *DpnII* and 80°C for 20 minutes for *HaeIII*.

Ethanol precipitation of digested products

In order to clean up the digested product from the unwanted reagents, an ethanol precipitation was preformed. For each sample, the reaction consisted of 125 µL of ice cold 95% ethanol, 1 µL 3 M sodium acetate (pH 4.6), and 1 µL of glycerol. After 30-minutes of incubation at 4°C, each sample was centrifuged at 3790 RPM for 30 minutes to pellet the precipitated DNA. After centrifugation, an additional 100 µL of ice cold 70% ethanol was added to each reaction and spun for an additional 30 minutes at 3790 RPM. Samples were then inverted to empty out the ethanol and centrifuged inverted at 700 RPM to remove excess ethanol and dry the pellet.

Capillary Gel Electrophoresis

To run the capillary gel electrophoresis, samples were first prepared by resuspending the pelleted DNA in 20 µL of a master mix containing 20 µL formamide and 0.25 µL CEQ 600 bp standard. A drop of mineral oil was placed in each sample to prevent evaporation. Once prepared, the samples were run using the Beckman Coulter CEQ 8800 genetic analysis system. Once the terminal restriction fragment peaks were produced the data was collected and analyzed.

Analysis of TRF data

The data obtained from the CEQ 8800 was analyzed using Principle Component Analysis (PCA). "Principle component analysis is a multivariate method that can be used to reduce the number of variables that need to be considered in an analysis and for interpretation of the data. Each principle component may be one of the original variables or else a linear combination of more than one of the original variables. The first principle component is identified as the single variable or linear combination of the original variables that has the maximum variance. The second principle component is the single variable or linear combination with the second highest amount of variance, subject to the condition that it is orthogonal to the first principle component. This orthogonality implies that the second principle component is independent of the first, which also means that it is not linearly correlated to the first."(3) The use of PCA in this study allowed us to reduce the large number of variables to just the ones that needed to be considered for the best interpretation of results.

Similarity between microbial communities was also assessed using cluster analysis. "The principle goal of cluster analysis is to classify samples or other objects into two or more 'natural' groupings based on similarities between the measurements that have been made on the samples." For this study Bray-Curtis similarity was used to determine the percent similarities among each of the samples. "The Bray-Curtis coefficient has been shown to offer a high level of power and robustness. Given these characteristics, the Bray-Curtis coefficient is ideally suited to T-RFLP data sets, whereas Euclidean methods are not." (4).

Results and Data

O_2 and TPH concentration found in samples

The results for the soil gas analysis of all the samples showed that oxygen content was higher soil near the willow compared to the control soil. In addition, as time progressed this difference appeared to increase. Meanwhile, soil analysis revealed that TPH concentrations, although varied, remains consistently lower in the willow samples compared to the control samples.

Figure1. Comparison of total petroleum hydrocarbon and oxygen content separated by depth and time. Samples were received in April, August, and December of 2004 and were renamed to spring, summer, and winter of 2004.

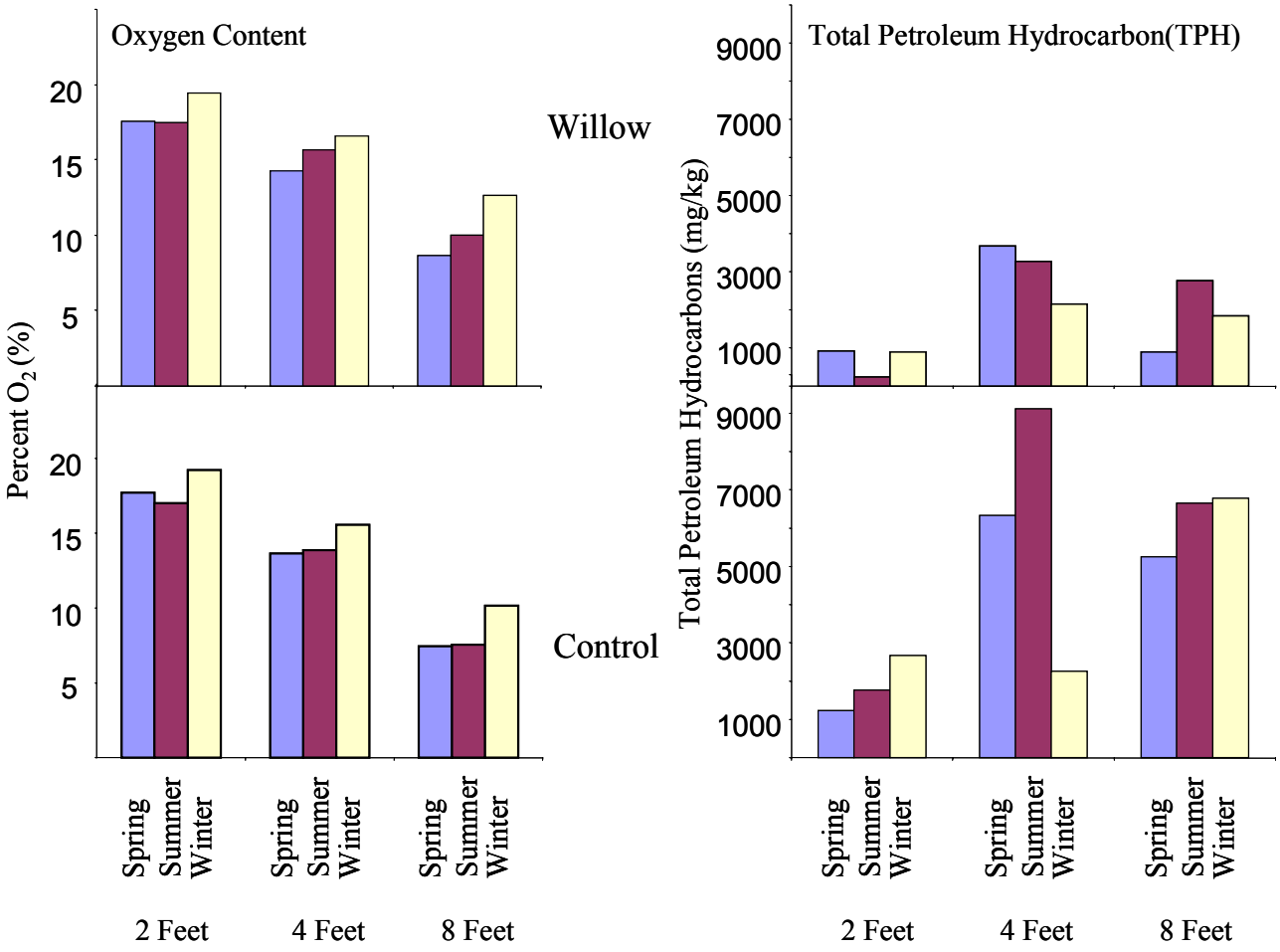
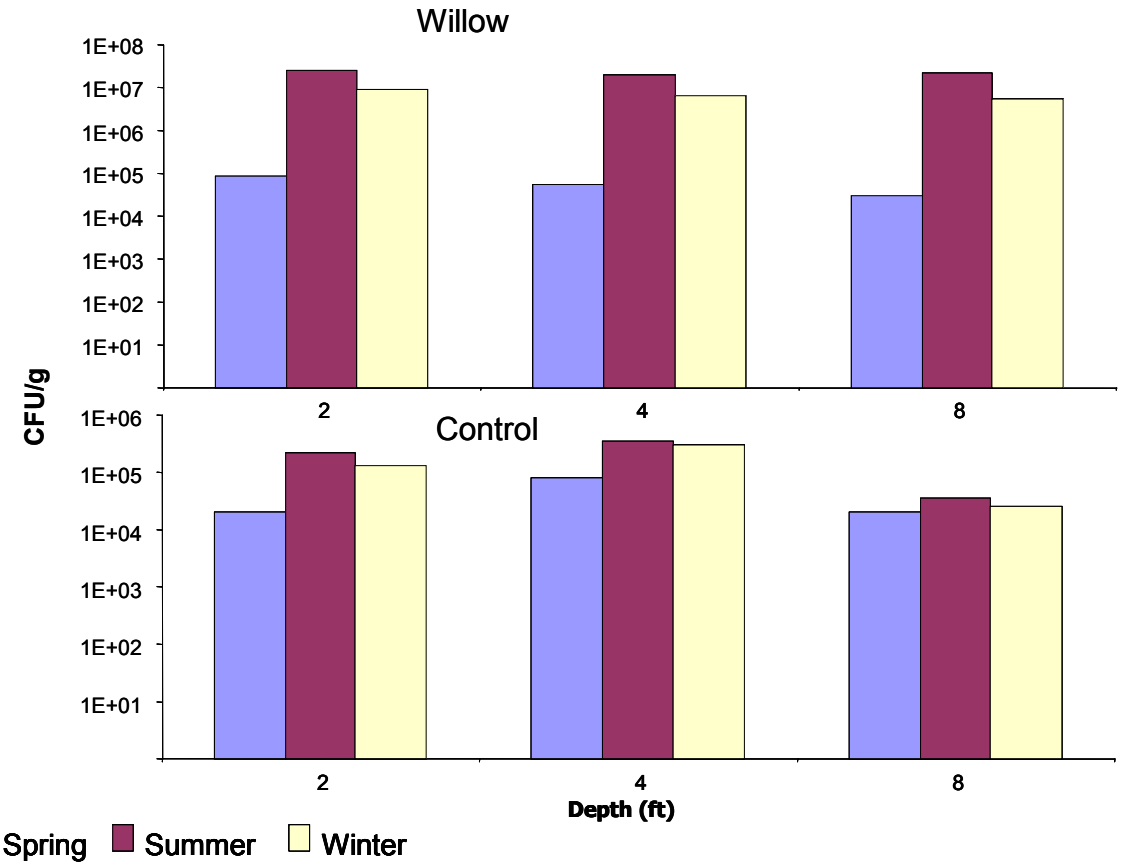


Plate counts

There was a large difference in the amount of aerobic heterotrophic bacteria found in the willow samples compared to the control samples. (Figure 2)

Figure 2. **H**eterotrophic aerobic bacteria found in each sample. Data was gathered and calculated after a 2 week incubation period on R2A agar @ 37°C. Willow samples on the top, control samples on the bottom.



Microbial Community Structure

PCA was performed with data from both Willow and Control TRF patterns created by digestion with both enzymes used. TRF data revealed that the microbial communities appeared to be changing over time (Figure 3, 4). Furthermore, the changes in community structure appeared to be related to the sampling time (season) and sample depth. The enzyme *Dpn* II did not yield as clear of results compared to the enzyme *Hae* III which gave a better separation by sampling time as well as depth.

Figure 3. PCA graphs representing the microbial communities for both control and willow samples. Digested with endonuclease *Dpn* II. The number after the sample name denotes the sampling depth. Blue lines indicate principle component loadings and represent the contribution of specific TRF peaks to the position of each sample in the PCA graph.

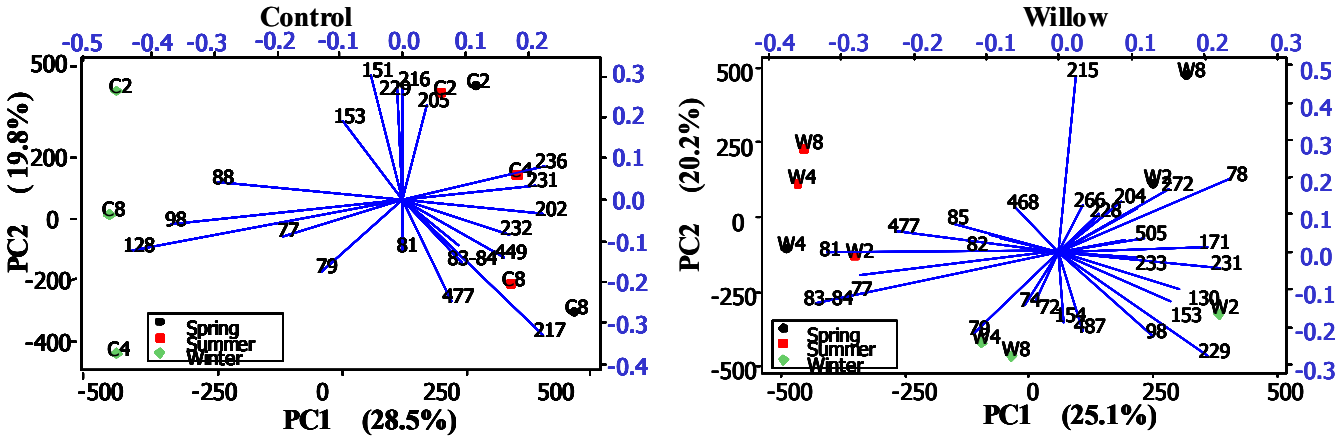
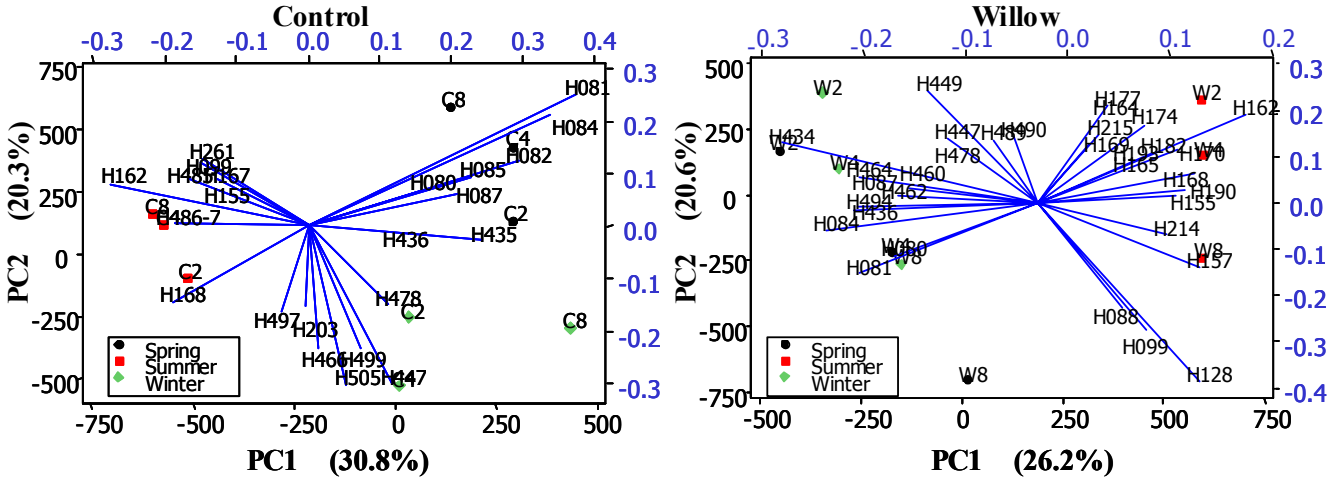


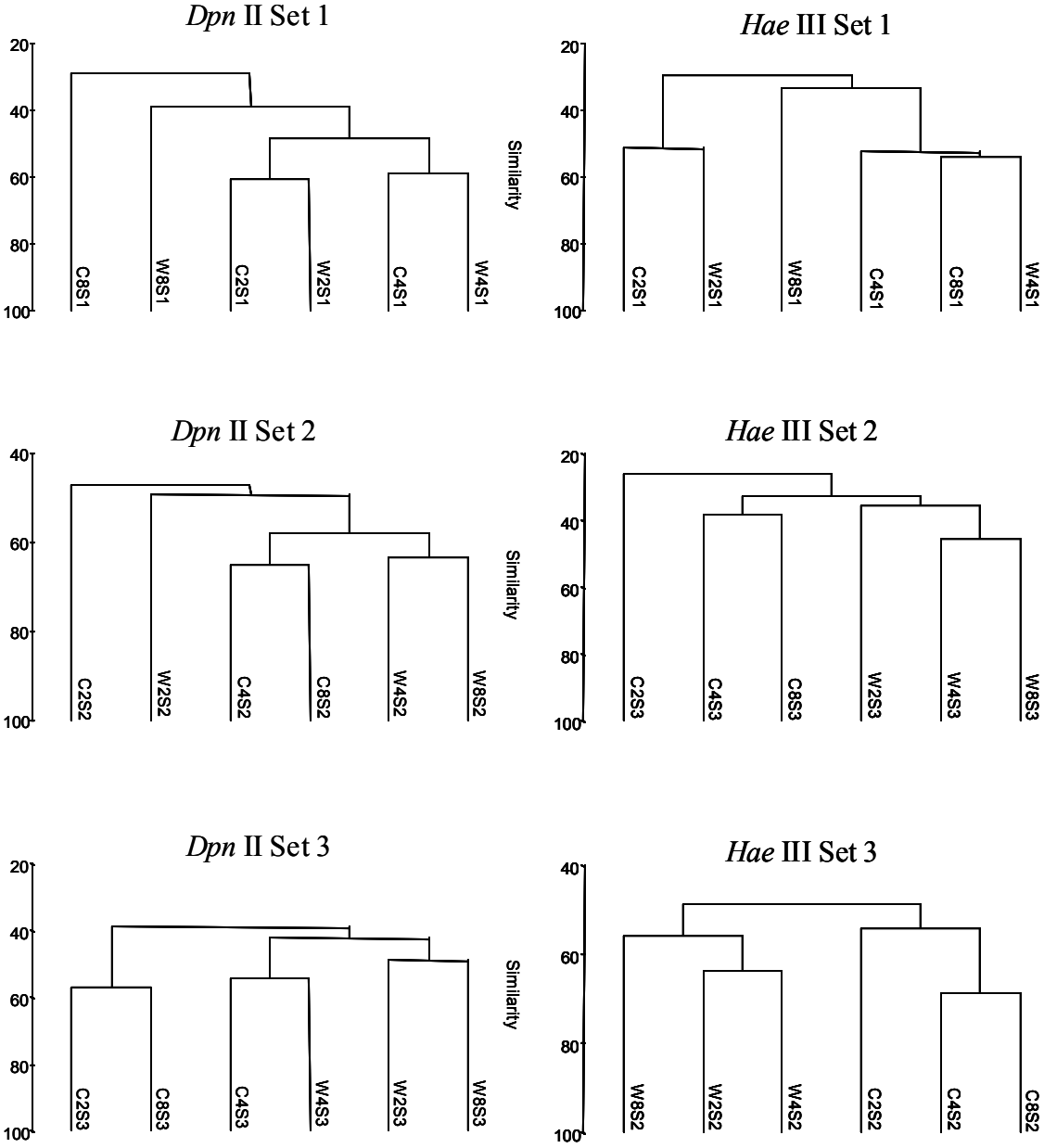
Figure 4. PCA graph representing the microbial communities for both control and willow samples. Digested with the endonuclease *Hae* III. The number after the sample name denotes the sampling depth. Blue lines indicate principle component loadings and represent the contribution of specific TRF peaks to the position of each sample in the PCA graph.



Similarity of microbial communities as determined by cluster analysis

Similarity cluster analysis showed that as time progressed, the microbial communities for the willow samples become more distinct and less similar to the microbial communities found in the control. (Figure 5). In both the willow and control samples the 2 foot sampling depth is the least similar to the other depths, indicating a possible role for oxygen levels in the structure of the microbial community.

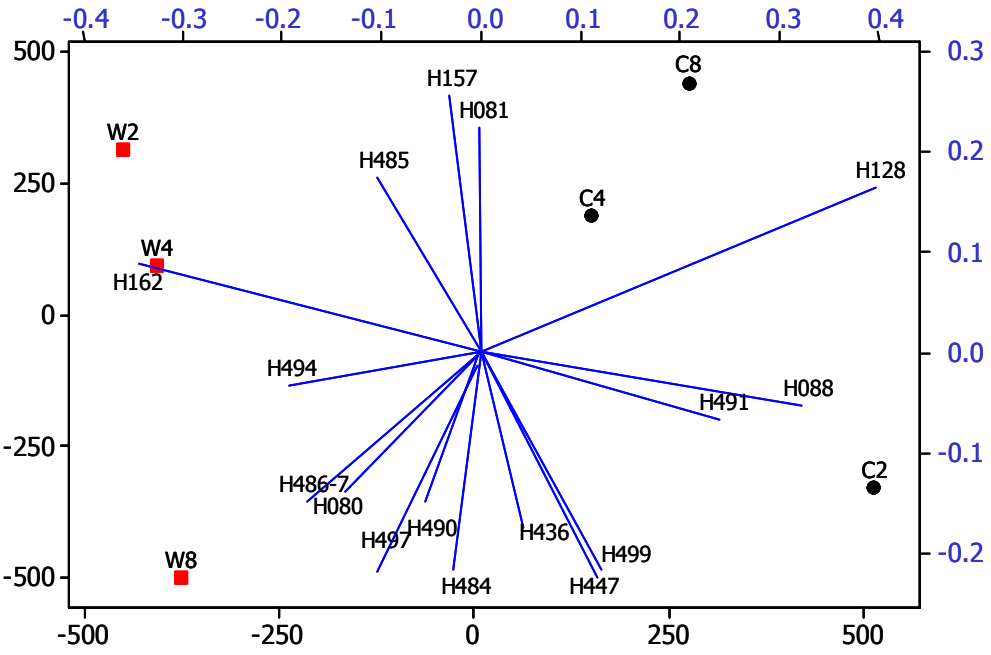
Figure 5. Bray-Curtis similarity charts comparing microbial community of willow and control soil. Set 1 samples were taken in April, Set 2 in August and Set 3 in December of 2004. 'W' denotes willow samples, 'C' denotes control samples. The sampling depth is indicated after the letter.



Identifying dominant TRF peaks

Compiling both control and willow samples into one PCA graph for the December sampling time(Figure 6) we were able to distinguish which peaks are most likely contributed to the differences between willow and control soils at the point where they were the least similar. The TRF data from the *Hae*III enzyme digest was used since it gave the best results. In ths graph there is a clear separation of willow soils from control soils. From the most prominent principle component loads (blue lines in Figure 6), the TRF peaks that contribute the most to differences between willow and control soils are peaks H485, H162, H486-7, H80, H491, H128, and H088.

Figure 6. PCA of Control and Willow soils during Winter 2004. The number after the sample name denotes the sampling depth. Blue lines indicate principle component loadings and represent the contribution of specific TRF peaks to the position of each sample in the PCA graph.



The types of bacteria represented by these TRF peaks are listed in Table 1. The most interesting identifications are from TRF peaks H485-7 which most likely represent *Pseudomonas* species; a genus well know for species with petroleum degrading abilities. It is also interesting to see that TRF peak H128, dominant in the lowest samples of the control soil, most likely represents anaerobic or facultative anaerobic bacteria.

Table 1. Tentative identification of organisms represented by TRF peaks from Figure 6.

HaeIII	
Peak #	**Possible organisms**
485	Flavobacterium salegens DSM 5424 (T). Flavobacterium lutescens ATCC 27951. Pseudomonas sp. 16S rRNA gene.
162	Desulfovibrio desulfuricans. Burkholderia sp. str. YY62. Azospirillum sp. Beta-proteobacterium species 16S rRNA gene (isolat Leptothrix sp. str. HS. Pseudomonas sp. 16S rRNA gene, isolate 150.
486	Pseudomonas B13 str. B13. Pseudomonas sp. 4FB7 16S ribosomal RNA gene, partial sequence Saprospira sp. SS91-40 gene for 16S rRNA, partial sequence
487	Pseudomonas stutzeri. Pseudomonas fluorescens 16S rRNA gene, strain CHA0 Uncultured beta proteobacterium clone GOBB3-CL103 16S ribosomal RNA gene, partial sequence
80	Streptomyces sp. 1038 gene for 16S rRNA, partial sequence
491	Pseudomonas stutzeri IAM 12668.
128	Fibrobacter succinogenes succinogenes (strain A3C) 16S ribosomal RNA Prevotella pallens 16S rRNA gene, strain 9423 Frankia sp. str. Ag45/Mut15.
88	Planctomyces sp. (strain 248) 16S rRNA gene, partial Microbacterium imperiale IFO 12610 (T).

Conclusions

According to aerobic heterotrophic plate counts, the willow tree seems to be promoting aerobic microbial growth in the soil. This coincides with data showing a higher oxygen content and lower TPH content in the willow soils. Examination of microbial community data showed that the sampling season also affects the composition of the microbial community as well as sampling depth (possibly related to soil oxygen content). Tentative identification of the microbial community members that contributed to differences between willow and control soils indicated a possible increase in the prevalence of aerobic bacteria with TPH degrading ability. Unfortunately, because TPH levels varied from sampling to sampling, it is not clear if any TPH was removed from the soil near the willow during this sampling period. However, lower TPH levels were consistently found in the willow samples compared to the control samples, which suggests that the willow tree and/or the associated bacteria are degrading the petroleum found in the soil. In order to be sure the willow tree is affecting TPH levels in the soil, more samples should be analyzed over a longer period of time.

Acknowledgements

I would like to thank everybody at the Environmental Biotechnology Institute for giving me the opportunity to work with them. I would also like to give special thanks to Dr. Chris Kitts for his continued support and knowledge, as well as Alice Hamrick for her support and guidance. Finally, I would like to thank Unocal for providing me with the project and a chance to make an impact on the environment.

References

1) U.S Environmental Protection Agency, 2004, Browfields Cleanup and Redevelopment. http://www.epa.gov/brownfields/about.htm

2) Interactice planning and management. 2001. Final Guadalupe restoration plan http://www.dfg.ca.gov/ospr/organizational/scientific/nrda/guadalupe.pdf

3) Hurst C., Crawford R., Knudsen G., McInerney M., Stetzenbach L. "Manual of environmental Microbiology" 2nd edition, ASM press, Washington D.C. Chapter 15 pages 166-177.

4) Gavin N. Rees, Darren S. Baldwin, Garth O. Wilson, Shane Perryman, Daryl L. Nielsen. 2004. Ordination and significance testing of microbial community composition derived from terminal restriction fragment length polymorphisms: application of multivariate statistics