

Coupling of Realistic Rate Estimates with Genomics for Assessing Contaminant Attenuation and Long-Term Plum Containment

Ronald L. Crawford
Environmental Biotechnology Institute
P.O. Box 441052
University of Idaho
Moscow, ID 83844-1052

This project is a collaborative project between the University of Idaho (Dr. Ronald L. Crawford, PI), the Idaho National Engineering and Environmental Laboratory (INEEL; Dr. Frederick Colwell, PI) and Northwind Environmental (Dr. Kent Sorenson, PI).

This report contains a summary of progress of only Dr. Crawford's part of the effort. This part of the overall project involves examination of the microbial genomics of the Snake River Plain Aquifer (SRPA) at the INEEL Test Area North (TAN) site. The work is being performed by two graduate students, Mr. Daniel Erwin and Ms. Amy Torguson. Progress is reported here as work done by each student. Graduate Student: Dan Erwin Objective. Mr. Erwin's primary objective is to create a non-culture based metagenomic library of the attached and planktonic microbial communities of the Snake River Plain Aquifer (SRPA). This library will then be probed for the presence of genes in the microbial communities that may be used in the cometabolism of chlorinated aliphatics such as trichloroethylene (TCE). Methods. Our approach to this objective includes four phases: sample collection, DNA processing, clone production, and library screening. Samples are collected for both attached and planktonic communities existing within the SRPA. The attached organisms are collected by suspending sterile substrate columns (basalt coupons) into the aquifer for a period of 3-6 months. Planktonic organisms are collected via hollow fiber filtration. Metagenomic DNA is obtained from the resulting biomass of each collection through a lysozyme/hot SDS treatment. Gel electrophoresis is then used to isolate 25-40 kb fragments of the randomly sheared DNA. These fragments are electroeluted, end repaired for ligation into Fosmid vectors, and used in the production of our metagenomic library. The libraries are then screened by macro-array analysis for genomic markers of such enzymes as methane monooxygenase and other alkane monooxygenases present in the indigenous microorganisms. Figure 1 (refer to separately mailed technical progress report) shows one of our field sampling campaigns in progress (water filtration) at the INEEL. Figure 2 (refer to separately mailed technical progress report) shows a close-up of the coupon apparatus used to contain basalt chips within the aquifer during growth of biofilms. Research status. During the first year of this project, efforts have been made to achieve three goals: community sampling, optimization of DNA extraction, and Fosmid clone production. The first goal has been met through the limited bias collection of planktonic as well as attached cellular communities. Both collections resulted in slurries consisting of significant amounts of inorganic impurities (basalt sediments). This has made it necessary to establish a method for genomic DNA isolation from these slurries for further processing. We have found that quality large fragment (25-40 kb) DNA from both the attached and planktonic microbial communities can be obtained by the direct cell lysis of recovered aquifer materials using a lysozyme/hot SDS treatment. However, the metagenomic DNA yield from each community using this method was found to be insufficient for cloning purposes. Since it appears that a primary limitation for production of our metagenomic library will be the quantity of high-quality DNA we can isolate from our environmental samples, we have examined a novel PCR procedure for amplifying total genomic DNA without bias. The method employs the REPLI-g Whole Genome Amplification (WGA) kit from Molecular Staging, Inc. Figure 3 (refer to separately mailed technical progress report) shows an example of our recent use of this technique to amplify SRPA microbial community DNA. We have evaluated this kit for non-biased amplification of all DNA in a sample for use in generating sufficient quantities necessary for building a metagenomic library, even from a very limited amount of community DNA. Real-time PCR utilizing primer sets designed for *pmoA* sequences specific to various methane-assimilating bacterial groups was used to test for bias in the system. Enrichments from both attached and unattached microbial communities were amplified by WGA and compared against unamplified samples. These experiments showed varying results between controls and amplified

environmental samples. In the case of the enriched samples, the amplified products showed little variation from the unamplified DNA (Fig. 4; refer to separately mailed technical progress report). However, the amplified environmental samples showed some discrepancies in the size of PCR amplicon produced (Fig. 5; refer to separately mailed technical progress report). Future research. Currently we are dealing with the issue of collecting sufficient quantities of genomic DNA from the well samplings. Additional substrate tubes have been placed into the well as the first set only resulted in a small amount of biomass. These tubes will be allowed to colonize for 8 months before removal and extraction of biomass, as before.

Metagenomic DNA will be obtained for each community through retrieval of the substrate tubes and filtration of 10,000 L of aquifer water. We believe that enough product can be acquired through previously stated methods to result in a yield sufficient for library production. While the REPLI-g amplification system remains an option, further studies are needed to determine the nature of the amplified product. Cloning of this amplified product has so far been unsuccessful and verification of its ability to amplify in a non-biased fashion is yet to be fully determined. Once an adequate amount of metagenomic DNA is acquired, it will be processed, as described above, to produce the clones necessary for our library. Probes used in the macro-array analysis of this library will be designed from alignments of all publicly available sequences for alkane monooxygenases listed in GenBank. Individual probes will be created based on substrate specificity and phylogenetic orientation. Graduate Student: Amy Torguson Objective. To characterize the microbial diversity associated with the planktonic and biofilm microbial populations present in a pristine deep-water well from the Snake River Plain Aquifer (SRPA), located at the Idaho National Engineering and Environmental Laboratory (INEEL) in Idaho Falls, Idaho. Methods. The Polymerase Chain Reaction (PCR), using two different sets of eubacterial primers, 27F-907R and 338F-907R, will be used to amplify the 16S rDNA from the microbial community present in the hollow-fiber filtered water sample collected from the SRPA at INEEL. The PCR product will then be cloned and transformed into *Escherichia coli*. Plasmid DNA will be purified and subjected to Restriction Fragment Length Polymorphism (RFLP) analysis using the enzyme *HaeIII* to determine the number of different microorganisms present in the water sample. Plasmid DNA with unique restriction patterns will be sequenced and analyzed using the Basic Local Alignment Search Tool (BLAST) and the Ribosomal Database Project (RDP) to determine the identity of the microorganisms. These analyses will determine if there are organisms indigenous to the SRPA known to possess the ability to degrade trichloroethylene (TCE). 16S rDNA analysis of both the planktonic and biofilm microbial communities will allow comparison of the two populations and thus determine where potential TCE degraders predominate. In addition, identification of the microorganisms present in the SRPA will aid in elucidation of the role that intrinsic bioremediation might play when chlorinated solvents such as perchloroethylene (PCE) and TCE unexpectedly contaminate a region of the aquifer. Research status. Because aquifer samples are valuable and difficult to obtain, we initially tested several methodologies for experimental condition optimization to ensure that our approach would work and water samples would not be wasted. Pure cultures of actinomycete isolates that are not easily amplified were used to optimize the PCR and Denaturing Gradient Gel Electrophoresis (DGGE) procedure. Four actinomycete isolates (high G + C Gram-positive bacteria) isolates were grown on yeast-dextrose agar and non-sporulating colonies were used for PCR. To extract the DNA necessary for PCR, individual colonies were added to 100 micro-L of TE buffer containing 1% Triton X-100 and boiled at 96 degrees C for 10 minutes to lyse the cells. The cell suspension was then cooled and centrifuged at 13,000 rpm for 2 minutes to pellet cell debris. The resulting supernatant was used in PCR reactions. Reaction conditions included a denaturation step at 95 degrees C for 5 min, followed by 30 cycles of denaturation at 95 degrees C for 1 min, annealing at 55 degrees C for 1 min, elongation at 72 degrees C for 2 min, and a final elongation step at 72 degrees C for 7 min. Two Eubacterial primer sets were used to amplify 16S rDNA genes: 27F-907R and 338F-907R. 27F: 5'-AGAGTTTGATCMTGGCTCAG-3' 338F: 5'-ACTCCTACGGGAGGCAGC-3' 907R: 5'-CCGTC AATTCMTTTRAGTTT-3' Three of the four actinomycete DNA samples amplified with the primer set 338F-907R. The amplified products were then run on a 40-80% DGGE gel. As expected from a pure culture, only single bands were obtained. To determine whether or not our primer sets and reaction conditions would work on Gram-negative organisms, several were grown in pure culture and mixed together to simulate a mixed population. A 100 micro-L sample of the mixed culture was centrifuged and lysed as described above. PCR was carried out according to the above conditions using both primer sets. Amplification products were obtained with both primer sets, and the products of each were run on a 40-80% DGGE gel. Multiple bands resulted from the Gram-negative mixed bacterial culture as expected; however, DGGE failed to resolve the expected number of bands, even when the denaturation gradient was

changed, indicating that DGGE would not be an appropriate choice for analysis of these potentially complex planktonic and attached microbial communities. Alternatively, RFLP analysis was chosen. Planktonic organisms were collected from the SRPA via hollow fiber filtration. Samples were concentrated and lysed as described above. DNA from the lysis procedure was used as template for 16S rDNA amplification using the primers 907R and 27F. The PCR product was purified (Mo Bio PCR cleanup kit) and cloned into pCR4-TOPO and transformed into Top10 E. coli competent cells (Invitrogen TOPO cloning kit). Nine hundred sixty transformants were picked into 96 well plates containing Hogness medium using the Qiagen BioPick robot. A duplicate archival plate was made of each of the 10 plates and stored at -80 degrees C. Plasmid isolations of 384 clones were performed using the Promega Wizard plasmid purification kit and the Qiagen BioRobot 3000. Restriction enzyme digests of the purified plasmid DNA was carried out in 30 micro-L reactions containing 1 micro-g plasmid DNA and 10 units of HaeIII for 1 hour at 37 degrees C. Digestions were analyzed on a 1.5% TAE agarose gel (Figure 6; refer to separately mailed technical progress report). Restriction patterns were compared for 29 clones. Of the 29 clones analyzed thus far, seven appeared to have distinct restriction patterns and were chosen for sequencing. Sequencing reactions were performed at the Washington State University Laboratory for Biotechnology and Bioanalysis in the forward and reverse direction using 907R and 27F as sequencing primers. Computations were performed at the Swiss Institute of Bioinformatics (SIB) using the BLAST network service; blosum 62 was the matrix used. All seven isolates had E values of 0.0. The results are summarized in Table 1 (refer to separately mailed technical progress report). DNA from attached communities, which were collected by suspending sterile substrate columns into the aquifer for 3-6 months, were also used as template for PCR of the 16S rDNA using 907R and 27F. Analysis of clones containing these amplicons is in progress. Future research. Analysis of the remaining 931 plasmid-containing ?planktonic? clones is currently in progress. As the restriction patterns are analyzed, the clones will be organized into groups with at least two clones sequenced from each group. We will analyze the attached microbial community structure by RFLP analysis and compare the emerging groups with those identified for the planktonic community. Algorithm Citation Altschul, S. F., T. L. Madden, A. A. Schaffer, J. Zhang, Z. Zhang, W. Miller, and D. J. Lipman. (1997), "Gapped BLAST and PSI-BLAST: a new generation of protein database search programs." *Nucleic Acids Res.* 25:3389-3402. Program Citation National Center for Biotechnology Information, Bldg 38A, NIH, 8600 Rockville Pike, Bethesda, MD 20894. Database Citation Benson, D.A., Boguski, M.S., Lipman, D.J., Ostell, J., and Ouellette, B.F.F. GenBank. *Nucl. Acids Res.* 26(1):1-7 (1998).