Analysis of Organisms Surviving in Highly Contaminated Environments

Carl Abulencia[1], Denise L. Wyborski[1], Bob Buchner[1], Cathy Chang[1], Joe Garcia[1], Terry C. Hazen[2], Keiko Obokata[1], Mircea Podar[1], Toby Richardson[1], Martin Keller[1]

[1] Diversa Corporation, San Diego, CA
[2] Lawrence Berkeley National Lab, Berkeley, CA

Diversa Corporation, 4955 Directors Place, San Diego, CA 92121, USA 858-526-5235; Fax 858-526-5735

Virtual Institute of Microbial Stress and Survival

Abstract

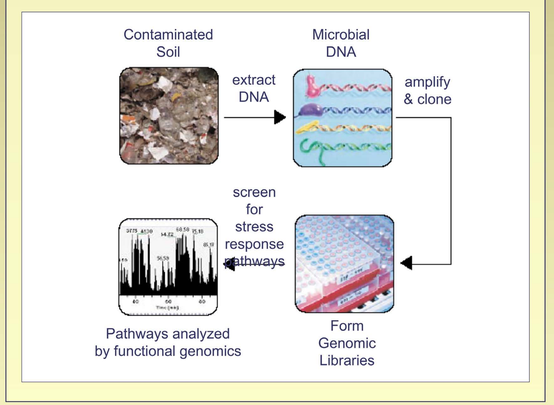
Background:
One of the main goals of the DOE is to clean up metal and radionuclide contamination of soil and groundwater at DOE sites. One possible solution to this problem is bioremediation. In order to understand bioremediation in contaminated environments, it is first necessary to examine the organisms present in these environments and analyze their responses to stress conditions on a genetic level.

Methods: We extracted high molecular weight DNA from organisms present in contaminated soil sediment samples using a method which preserves the integrity of the DNA. Because the number of organisms in these samples was low, the genomic DNA was amplified using a phage polymerase amplification system. 16S rRNA analysis was then used to examine the microbial diversity of the samples. The amplified DNA was also used in the construction of large and small insert DNA libraries. These libraries were then screened for the presence of histidine kinase genes with homology to a subfamily of *Desulfovibrio vulgaris* histidine kinases.

Results: Genomic DNA has been extracted and amplified from nine different sites at the NABIR field research center. 16S rRNA analysis revealed the presence of distinct bacterial phyla, including proteobacteria, acidobacteria, and planctomycetes. Small and large insert libraries were constructed for all samples and examined for clonal diversity. Plaque hybridization of these libraries to histidine kinase homologous probes resulted in multiple positive clones. These clones will be compared and used to develop a better understanding of cellular responses to different environmental factors.

Conclusion: These experiments have furthered the understanding of how the biological organisms in a contaminated system are organized, regulated and linked. This will ultimately lead to an understanding of the ability of these organisms to attenuate metal and radionuclide contamination, and help to provide a method for maximizing this attenuation.

Technology Flow For Genomes To Life Program



Sample Data

NABIR FRC Field Sampling Plan

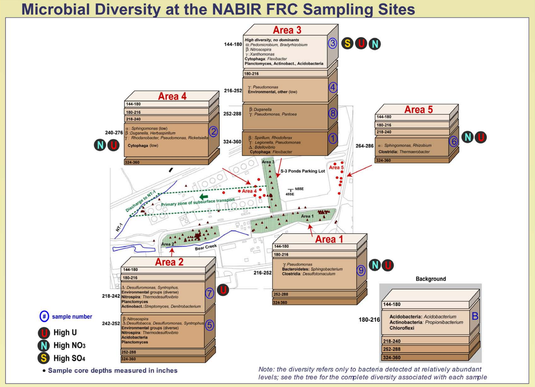
Soil core samples were obtained from the following areas:

Background site: 2 samples at least 1 m (~39 inches) depth.
Area 1: high Uranium (U) (>1mg/l), high NO_3 (>1000mg/l), low pH (3.25-6.5).
Area 2: high U, low NO_3 (<100mg/l), neutral pH, higher dissolved oxygen
Area 3: highest U, highest NO_3, high SO_4 (1000mg/l), low pH (<4.0)
Area 4: high U, high NO_3, neutral pH
Area 5: high U, high NO_3, neutral pH

Large insert DNA extractions were performed on these samples.

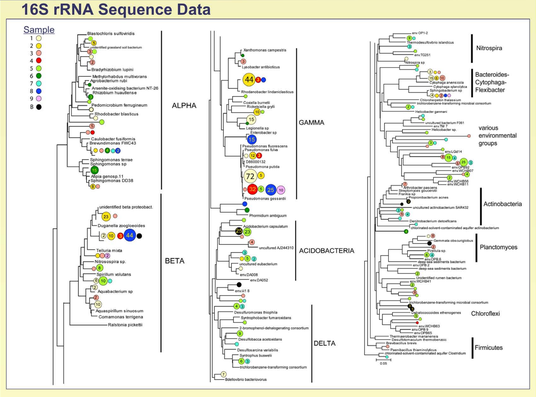
Results

Microbial Diversity at the NABIR FRC Sampling Sites



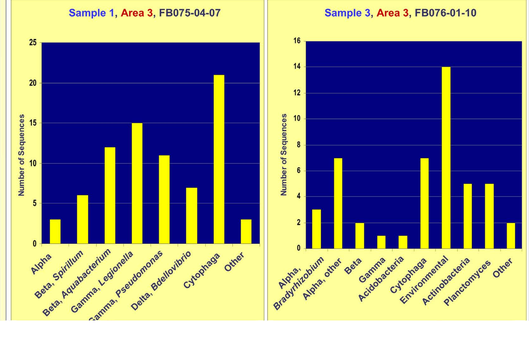
Results

16S rRNA Sequence Data



Results

16S rRNA Sequence Analysis



Results

Library Clone Sequencing Data

Library 3868, Sample 1, Area 3, FB-075-04-07

Random clones from the libraries were end-sequenced. The following are a sample of Blast hits from these clones.

Total Clone-ends Sequenced	1270	
Identical Clones	52	4.1%
Clones with no Blast hits	155	12.2%

E value	Top Blast hit
3.00E-59	gi|47059343| heavy metal transporting ATPase [Ornithobacterium rhinotracheale]
3.00E-46	gi|39996434| heavy metal efflux pump, CzcA family [Geobacter sulfurreducens PCA]
1.00E-17	gi|48846102| COG1235: Metal-dependent hydrolases of the beta-lactamase superfamily I [Geobacter metallireducens GS-15]
3.00E-38	gi|29611404| FeS-binding reductase-like protein [uncultured bacterium]
1.00E-45	gi|53760196| COG0347: Nitrogen regulatory protein PII [Methylobacillus flagellatus KT]
3.00E-24	gi|46131534| COG0715: ABC-type nitrate/sulfonate/bicarbonate transport systems, periplasmic components [Ralstonia eutropha JMP134]
1.00E-10	gi|62484831| predicted cytotoxic translational repressor of toxic-antitoxic stability system [uncultured bacterium]
2.00E-56	gi|46579304| aminotransferase, classes I and II [Desulfovibrio vulgaris subsp. vulgaris str. Hildenborough]
2.00E-67	gi|48733601| COG0642: Signal transduction histidine kinase [Pseudomonas fluorescens PfO-1]
5.00E-19	gi|48783463| COG4564: Signal transduction histidine kinase [Burkholderia fungorum LB400]
8.00E-21	gi|34104565| probable sensory transduction histidine kinase [Chromobacterium violaceum ATCC 12472]
4.00E-21	gi|52003208| two-component sensor histidine kinase [Bacillus licheniformis ATCC 14580]
5.00E-21	gi|46164990| COG0642: Signal transduction histidine kinase [Pseudomonas aeruginosa UCBPP-PA14]

Results

Library Clone Sequencing Data

Library 3875, Sample 3, Area 3, FB-076-01-10

Random clones from the libraries were end-sequenced. The following are a sample of Blast hits from these clones.

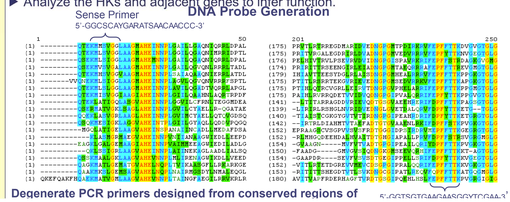
Total Clone-ends Sequenced	1119	
Identical Clones	8	0.7%
Clones with no Blast hits	206	16%

Evalue	Top Blast hit
1.00E-11	gi|15977112|ref|NP_578469.1| heavy-metal transporting cpx-type atpase [Pyrococcus furiosus DSM 3638]
2.00E-87	gi|56460133|ref|YP_155414.1| Predicted extracellular metal-dependent peptidase [Idiomarina loihiensis L2TR]
2.00E-62	gi|47571750|ref|NP_00241799.1| COG1230: Co/Zn/Cd efflux system component [Rubrivivax gelatinosus PM1]
7.00E-27	gi|39995602|ref|NP_951553.1| iron-sulfur cluster-binding protein [Geobacter sulfurreducens PCA]
7.00E-07	gi|46316764|ref|ZP_00217343.1| COG1561: Uncharacterized stress-induced protein [Burkholderia cepacia R18194]
2.00E-11	gi|53685719|ref|ZP_00099333.2| COG5557: Polysulphide reductase [Desulfitobacterium hafniense DCB-2]
1.00E-85	gi|48855283|ref|ZP_00309442.1| COG0820: Predicted Fe-S-cluster redox enzyme [Cytophaga hutchinsonii]
9.00E-16	gi|46580259|ref|YP_011067.1| CBS domain protein [Desulfovibrio vulgaris subsp. vulgaris str. Hildenborough]
2.00E-12	gi|48854666|ref|ZP_00308827.1| COG0642: Signal transduction histidine kinase [Cytophaga hutchinsonii]
7.00E-20	gi|23012938|ref|ZP_00052912.1| COG0642: Signal transduction histidine kinase [Magnetospirillum magnetotacticumMS-1]
1.00E-12	gi|46113064|ref|ZP_00200680.1| COG0642: Signal transduction histidine kinase [Exiguobacterium sp. 255-15]
5.00E-44	gi|13476679|ref|NP_108248.1| two-component sensor histidine kinase [Mesorhizobium loti MAFF303099]
2.00E-45	gi|37523305|ref|NP_926682.1| two-component sensor histidine kinase [Gloeobacter violaceus PCC 7421]
1.00E-12	gi|46113064|ref|ZP_00200680.1| COG0642: Signal transduction histidine kinase [Exiguobacterium sp. 255-15]

Methods

Library Screening

- ► Histidine Kinase protein superfamily chosen for sequence-based discovery.
- ► HKs contain phosphotransfer-mediated signaling pathways that allow cells to sense and respond to environmental stimuli.
- ► HK primers designed based on *Desulfovibrio vulgaris*, which is a metal and radionuclide reducing bacterium.
- ► Retrieve histidine kinase and response regulator sequences from other Proteobacteria in environmental samples by DNA hybridization using *D.vulgaris* HK probes.
- ► Analyze the HKs and adjacent genes to infer function.



Results

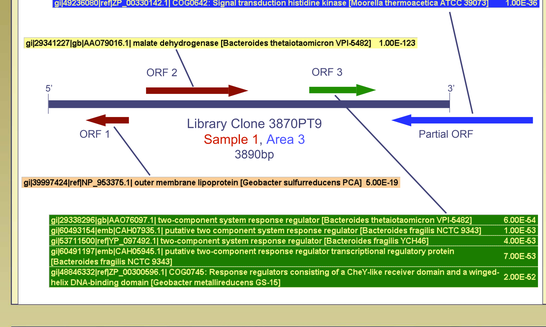
Histidine Kinase Genes From the Environment

Screening Progress

Sequence-Based Hybridization	26 clones in sequencing
Random Clone Sequencing	34 clones fully sequenced

Sequence analysis of a library clone shows a histidine kinase, response regulator, and nearby genes.

gi|29338296|gb|AAO76096.1| two-component system sensor histidine kinase [Bacteroides thetaiotaomicron VPI-5482]	8.00E-46
gi|60491198|emb|CAH05945.1| putative two-component regulator sensor kinase [Bacteroides fragilis NCTC 9343]	6.00E-44
gi|53711501|ref|YP_097493.1| two-component system sensor histidine kinase [Bacteroides fragilis YCH46]	8.00E-44
gi|29337986|gb|AAO76788.1| two-component system sensor histidine kinase [Bacteroides thetaiotaomicron VPI-5482]	2.00E-43
gi|53713476|ref|YP_099467.1| two-component system sensor histidine kinase [Bacteroides fragilis YCH46]	7.00E-42
gi|60493165|emb|CAH07835.1| putative two component system sensor histidine kinase [Bacteroides fragilis NCTC 9343]	9.00E-42
gi|49238690|ref|ZP_00330142.1| COG0642: Signal transduction histidine kinase [Moorella thermoacetica ATCC 39073]	1.00E-36



gi|29338296|gb|AAO76097.1| two-component system response regulator [Bacteroides thetaiotaomicron VPI-5482]	6.00E-54
gi|60493154|emb|CAH07936.1| putative two component system response regulator [Bacteroides fragilis NCTC 9343]	1.00E-53
gi|53711500|ref|YP_097492.1| two-component system response regulator [Bacteroides fragilis YCH46]	4.00E-53
gi|60491197|emb|CAH05945.1| putative two-component response regulator transcriptional regulatory protein [Bacteroides fragilis NCTC 9343]	7.00E-53
gi|48645332|ref|ZP_00300596.1| COG0745: Response regulators consisting of a CheY-like receiver domain and a winged-helix DNA-binding domain [Geobacter metallireducens GS-15]	2.00E-52

Conclusions

- ► **One of the ultimate goals of the DOE Genomics: GTL program is to understand what is involved in a microorganism's ability to survive in contaminated environments while reducing metals and radionuclides. The preceding work describes methods for accessing DNA from microorganisms found at contaminated sites, amplifying this genomic DNA, constructing libraries, and screening these libraries for genes involved in stress response using a combination of PCR, plaque hybridization, and library clone sequencing.**
- ► **Minute quantities of gDNA have been extracted from nine different environmental soil core samples.**
- ► **An optimized protocol for Multiple Displacement Amplification has been employed to obtain enough gDNA from these samples for library construction.**
- ► **16S rRNA sequences from these genomic samples have been analyzed to fingerprint the microbial community from which the samples were obtained, and to examine relationships between sample site and organisms found.**
- ► **16S rRNA analysis and library clone sequences from Area 3, which had the highest level of contaminants, showed a high level of diversity.**
- ► **End sequencing of small insert library clones indicated a large sequence diversity and resulted in 34 unique confirmed inserts with histidine kinase homology. Preliminary PCR screening and plaque hybridization resulted in 26 unique HK clones.**

Acknowledgements

We would like to acknowledge the following for their contribution to this work:

Diversa's Sequencing group for their phenomenal sequencing effort.

Terry Hazen's lab at LBNL, especially Eoin Brodie, Sharon Borglin, and Dominique Joyner for their help with sample handling and for the use of their radioactive containment lab.

Tamas Torok at LBNL for his help with sample handling.

David Watson at ORNL for sample collection.

DOE Genomics: GTL program for funding.